2014

Diagnostic Procedures for Preschoolers with Attention Deficit Hyperactivity Disorder with and without Comorbid Oppositional Defiant Disorder: A Comparison of Dimensional and Categorical Approaches to Classification

Lisa DeKett
Lehigh University

Diagnostic Procedures for Preschoolers with Attention Deficit Hyperactivity Disorder with and without Comorbid Oppositional Defiant Disorder: A Comparison of Dimensional and Categorical Approaches to Classification

by

Lisa Winders DeKett

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

School Psychology

Lehigh University

May 2014

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Date

_____
Patricia Manz, Ph.D.
Dissertation Advisor

_____
Accepted Date

Committee Members

_____
George DuPaul, Ph.D.
Committee Member

_____
Christine Cole, Ph.D.
Committee Member

_____
Stephen Soffer, Ph.D.
Committee Member

TABLE OF CONTENTS

List of Tables

List of Figures

Abstract

The early onset, temporal stability, and adverse outcomes associated with neurodevelopmental and disruptive behavior disorders underscore the importance of early identification and treatment. However, there are significant obstacles to diagnosis at the preschool level. This study examined the ways in which dimensional measures of behavior (parent ratings, teacher ratings, and direct observation) differentiated among preschoolers who met DISC-IV criteria for ADHD and/or ODD. Cluster analysis suggested the presence of two distinct clusters. The first cluster (*Significant Concerns*) was characterized by more elevated concerns in every area in which significant between-cluster differences occurred (oppositional/delinquent behavior, social problems, hyperactivity, thought problems, anxious/shy behavior, and perfectionism). The second cluster (*Moderate Concerns*) exhibited more inattention and problem behaviors than typical peers (per inclusion criteria), but were rated as exhibiting lower levels of problem behavior and social problems than Cluster 1 participants. Teacher ratings also reflected stronger social skills among Cluster 2 participants. Chi square analysis indicated no significant relationship between cluster membership and diagnostic classification. The diagnostic composition of both clusters was highly similar. The results indicated that the clusters were distinguished primarily by differences in symptom severity, rather than symptom presence/ absence. Parent and teacher ratings followed a similar pattern, being more elevated for Cluster 1 than Cluster 2 on nearly every variable. Implications for assessment and treatment are considered and possible directions for future research are suggested.

Diagnostic Procedures for Preschoolers with Attention Deficit Hyperactivity Disorder with and without Comorbid Oppositional Defiant Disorder: A Comparison of Dimensional and Categorical Approaches to Classification

Chapter I: Introduction

The prevalence of and adverse outcomes associated with neurodevelopmental disorders (NDs) and disruptive behavior disorders (DBDs) are well-documented and underscore the importance of early identification and intervention. Although a great deal of research has focused on identification of school-aged children at risk for NDs and/or DBDs, a growing literature base supports the conclusion that problem behaviors emerge at a very early age (Gimpel & Kuhn, 2000). For a significant number of preschool-aged children, the problem behaviors are of sufficient severity to meet diagnostic criteria (American Academy of Pediatrics, 2011; Lavigne et al., 1996). Studies examining the occurrence of psychiatric disorders within the preschool-aged population indicate rates similar to those reported among school-aged children (Egger & Angold, 2006; Lavigne et al., 1996). The prevalence of psychiatric disorders is even higher among clinic-referred preschoolers (Keenan & Wakschlag, 2000; Wilens, Biederman, Brown, Monuteaux, Prince & Spencer, 2002).

The presence of significant problem behaviors among preschool-aged children is of consequence because rather than reflecting a developmental phase, evidence indicates that early-onset behavior problems tend to persist over time (Campbell, Breaux, Ewing & Szumowski, 1984; Keenan et al., 2011), including periods of up to 10 years (Pierce, Ewing & Campbell, 1999). In addition to their temporal stability, early problem behaviors also warrant concern due to their association with later psychiatric disorders (Speltz, McClellan, DeKlyen & Jones, 1999). The link between early problem behaviors and subsequent diagnosis has been found for

disruptive and aggressive behaviors as well as behaviors characteristic of Attention Deficit Hyperactivity Disorder (ADHD; Burns & Walsh, 2002; Harvey, Youngwirth, Thakar & Errazuriz, 2009; Lahey et al., 2004; Lee, Lahey, Owens & Hinshaw, 2008).

Children with comorbid disorders appear to be at particular risk for long-term stability of symptoms and adverse outcomes (Speltz et al., 1999). Comorbidity is associated not only with problem stability, but also with increased number and severity of symptoms, both among school-aged children (Abikoff et al., 2002; Biederman et al., 2006; Jensen et al., 2001; Newcorn et al., 2001) and among preschoolers (Cunningham & Boyle, 2002; Gadow & Nolan, 2002). A meta-analysis of studies examining the developmental trajectory of various symptoms revealed that children manifesting both conduct problems and symptoms of ADHD exhibited a broader range of problem behaviors and more severe conduct problems at follow-up than children who exhibited either category of symptoms alone (Waschbusch, 2002).

Overwhelmingly, the evidence indicates that comorbidity is the rule, rather than the exception, among preschoolers with an ND or DBD (Kaplan, Dewey, Crawford & Wilson, 2001; Wilens, Biederman, Brown, Monuteaux, Prince & Spencer, 2002). Estimates suggest that as many as 60-70% of preschoolers with ADHD have comorbid Oppositional Defiant Disorder (ODD) (Kadesjo, Hagglof, Kadesjo & Gillberg, 2003; Keenan & Wakschlag, 2000). The high rate of psychiatric comorbidity found among preschoolers most commonly consists of dual diagnoses of ADHD and ODD.

Although the prevalence of NDs and DBDs among preschoolers suggests an extant structure for early diagnosis, the nature of young children's behavior is one among numerous obstacles to effective diagnosis at the preschool level. Many of the behaviors that characterize NDs and DBDs are, to some extent, common among preschoolers (Bussing, Lehninger & Eyberg, 2006;

Connor, 2002; Keenan, Shaw, Walsh, Delliquadri & Giovannelli, 1997; Keenan & Wakschlag, 2000). Short attention span, hyperactivity, noncompliance, impulsivity, and even aggression are developmentally typical for preschool-aged children (Byrne, Bawden, Beattie & DeWolfe, 2000; Connor, 2002; Keenan & Wakschlag, 2000; Keenan et al., 1997). For this reason, it is the frequency, severity, and persistence of such characteristics (rather than their mere presence) that suggest the possible presence of a psychiatric disorder (Carter, Briggs-Gowan & Davis, 2004; Kadesjo et al., 2003).

The situational specificity of behavior, clearly evident at the preschool level (Perrin, 2000), also complicates the diagnostic process (Carter et al., 2004). Time of day, activity type, and the varying expectations of different adults all contribute to fluctuations in young children's behavior across settings (Brandau & Pretis, 2004). Moreover, unlike school-aged children, younger children do not necessarily receive comparable early educational experiences.

Not all children attend preschool; among those who do, variability in the nature and quality of the preschool setting can impact symptom manifestation. Various symptoms, such as inattention, may not become evident until children encounter the cognitive, behavioral, and social demands of the school environment (Arons, Katz-Leavy, Wittig & Holden, 2002; Byrne et al., 2000). Although many young children attend preschool or day care, the structure and activities of these settings are qualitatively different from those of an academic setting (McGoey, DuPaul, Eckert, Volpe & Van Brakle, 2005). Lack of exposure to a setting that requires focused attention, persistence to task, compliance with behavioral demands, and inhibition of impulses may preclude the appearance of behaviors characteristic of NDs and/or DBDs.

Variations in exposure to an early educational environment impact not only symptom manifestation, but also the diagnostic process itself. Lack of consistent preschool attendance

limits the sources of information that can be utilized in assessment. It also precludes clinicians' capacity to ascertain the presence of impairment in multiple settings (American Academy of Pediatrics, 2011; Gimpel & Kuhn, 2000).

Such obstacles to effective identification and assessment of symptoms raise important questions regarding the utility of the existing diagnostic criteria when applied to the preschool population. The behaviors identified as symptoms of various disorders were designed for application within the upper childhood through adult age range with regard to both nature and threshold (Brandau & Pretis, 2004; Egger & Angold, 2006; McGoey et al., 2005; Thorell & Wählstedt, 2006). Preschool-aged children may be developmentally incapable of exhibiting some of the behaviors currently included among the diagnostic criteria for various NDs and/or DBDs (Keenan et al., 2011; Keenan & Wakschlag, 2002). In addition, rapid developmental changes may prevent children from meeting diagnostic criteria requiring that symptoms persist for a specified period of time.

The requirement of cross-situational impairment as an essential component for diagnosis has also been subject to criticism. As noted previously, this criterion is difficult to establish for children who are not exposed to an early educational setting. In addition, research findings indicate that a substantial number of children who were rated as impaired in only one setting exhibited an elevated level of functional impairment and later met full criteria for ADHD (Lahey et al., 2004; Lahey, Pelham, Loney, Lee & Willcutt, 2005). These findings suggest that the manifestation of symptoms across settings may not be a valid requirement for some preschool diagnoses.

Even more fundamental than criticisms of the content, nature, and relevance of the current diagnostic criteria are concerns regarding the underlying nature of the diagnostic process. The

Diagnostic and Statistical Manual of Mental Disorders (DSM, American Psychiatric Association, 2000) is based on a categorical approach, in which disorders are conceptualized as falling into separate and distinct categories (Achenbach, Becker, Döpfner, Heiervang, Roessner, Steinhausen & Rothenberger, 2008; Burns, Walsh, Owen & Snell, 1997), formulated through expert review and consensus (Achenbach, 2001). The categorical method is based on nosological models (Achenbach et al., 2008); diagnostic criteria are selected on the basis of their ability to predict the presence of a disorder (Burns et al., 1997). Under the categorical approach, diagnostic decisions are based on the presence or absence of symptoms as well as such criteria as age of onset, duration of symptoms, and functional impairment (Biederman et al., 1993). Potential advantages of a categorical approach to diagnosis include clinical utility, direct application to the disease model of psychiatry (Fergusson & Horwood, 1995), and improved communication among professionals.

Despite its status as the most widely used method of diagnostic classification, the categorical approach has been criticized at both the theoretical and the practical level. The traditional, categorical approach to diagnosis has been described as a "top-down" approach, originating with a theoretical model of a disorder that leads to identification of the symptoms required for diagnosis (Achenbach, Dumenci & Rescorla, 2003; Kasius, Ferdinand, van den Berg & Verhulst, 1997). Kasius and colleagues (1997) asserted that such an approach can result in a relatively arbitrary set of diagnostic criteria.

The absence of guidelines for determining the presence or impact of symptoms poses an obstacle at the practical level (Achenbach, 2001; Postert, Averbeck-Holocher, Beyer, Müller & Furniss, 2009). The reliability of diagnostic decisions is limited by such factors as individual differences in clinicians' training and skills (Cluett et. Al., 1998) as well as differences in

practitioners' experience and inductive reasoning processes (Jensen, Salzberg, Richters & Watanabe, 1993). Reliability is also limited by the fact that diagnoses are frequently based on information provided by a single informant (typically the parent) in a clinical setting (Cluett et al., 1998; Keenan et al., 2011), often in the context of a single session.

In light of the limitations of a categorical, theory-based method of diagnosis, an alternative, "dimensional" approach has been proposed. This approach conceptualizes "symptoms" as falling along a continuum of severity, rather than existing as distinct risk factors (Sonuga-Barke, Auerbach, Campbell, Daley & Thompson, 2005). Identification of a disorder via the dimensional approach is typically accomplished through use of rating scales that assess not only the presence, but also the frequency and/or severity of symptoms. Cut-off points are used to differentiate typical from deviant behavior, based on data regarding the frequency and intensity of various behaviors within a normative sample (Jensen et al., 1996). Elevated ratings (i.e., above designated cut-points) are considered risk factors for a disorder, rather than clear-cut evidence that a diagnosis is warranted (Egger & Angold, 2006).

Although the diagnostic process ultimately requires a categorical decision regarding the presence or absence of a psychiatric disorder, the distinctions between the categorical and dimensional approaches lie in the type(s) of data gathered and the method(s) of data collection. A categorical approach relies on information about the number of symptoms present, their time of onset, and their chronicity; this information is typically gathered through parent interview. Diagnosis via a dimensional approach typically involves gathering data from several informants (e.g., parents, teachers, caregivers) through rating scales/questionnaires (Achenbach et al., 2008). Dimensionally-based assessment might also include behavioral observation to allow for more direct assessment of the child's behavior within the context of the natural environment. The

increased structure and specificity of the data collection methods and the inclusion of multiple sources of information that characterize the dimensional approach are likely to improve the reliability and validity of the resulting diagnostic decisions, both within and across practitioners.

Although there are numerous advantages to dimensional diagnosis, this approach also entails several limitations, including its reliance on behavior rating scales. Many behavioral rating scales have been criticized due to lack of specificity and precision in response options (Egger & Angold, 2006) as well as inconsistency across respondents (Achenbach et al., 2008; Carter et al., 2004). Concerns have also been expressed regarding the utility of behavior rating scales for behaviors or disorders with low prevalence rates (Edelbrock & Costello, 1988). Questionnaires frequently do not assess the presence of rare behaviors/symptoms, as their inclusion would likely have an adverse impact on the psychometric properties of the rating scale (Carter et al., 2004; Postert et al., 2009).

Another obstacle to the effective use of rating scales pertains to their interpretation. Selection of a cut-off score distinguishing normative from atypical levels of behavior has been described as a subjective determination requiring a compromise among various factors, including sensitivity, specificity, positive predictive power (the degree to which presence of a symptom predicts presence of a disorder), and negative predictive power (the degree to which absence of a symptom predicts absence of a disorder) (Carter et al., 2004; Chen, Faraone, Biederman & Tsuang, 1994; Hudziak, Copeland, Stanger & Wadsworth, 2004). Establishment of a cut-off score has significant implications, given that a small change in the raw score on a rating scale can result in substantial differences in the number and identity of individuals identified as symptomatic (Serna et al., 2002). Rating scales are commonly designed to identify children who fall above the 95th percentile; however, this approach may be unacceptable for clinical use, given

the potential for under-identification (Hudziak et al., 2004).

Practical utility can also be considered a disadvantage of the dimensional approach to diagnosis. Gathering information from multiple sources can be time-consuming. In addition, for many practitioners, direct observation may not be a feasible component of assessment.

The primary theoretical disadvantage of a dimensional approach to diagnosis pertains to the underlying conceptualization of the nature of psychiatric disorders. Although the dimensional approach has theoretical and practical relevance as well as empirical support, elevation in rating scale scores may be "neither a necessary nor sufficient condition for identification of a psychiatric disorder" (Egger and Angold, 2006, p. 317). Disorders have traditionally been conceptualized in terms of a medical model, in which diagnoses are considered to reflect a complex, underlying construct that is characterized by various features that may not be fully captured by rating scale measures. Reliance on an empirical method of identifying symptoms that fall above normative levels may not provide an evaluation that is sufficiently comprehensive, given the multifaceted nature of psychiatric disorders.

In many ways, the dimensional and categorical approaches to diagnosis appear to share an inverse relationship, with the strengths of one approach constituting the limitations of the other. Systematic analysis and comparison of the two methods would be beneficial in terms of providing further insight into their relative strengths and weaknesses, as well as the degree to which they produce similar or divergent outcome decisions regarding the presence or absence of a disorder. Examination of both dimensional and categorical diagnostic procedures could also indicate potential ways in which the two methods might be used in tandem (Biederman et al., 1993; Edelbrock & Costello, 1988; Speece, 1994-1995).

The current knowledge base regarding different methods of diagnosis and their relevance for

various populations is lacking in a number of areas. Limited information is available regarding the potential role of alternative approaches to the traditional, categorical method of identification (e.g., dimensional measures). The existing research base is particularly limited with regard to diagnosis at the preschool level. The nature of the existing diagnostic criteria may be less appropriate for this age range, given uncertainty regarding the distinction between developmentally appropriate versus maladaptive levels of behavior (Brandau & Pretis, 2004; Byrne et al., 2000; Campbell et al., 1984; Carter et al., 2004; Gimpel & Kuhn, 2000) as well as questions about whether preschool-aged children are capable of engaging in some of the purposeful behaviors that characterize such disorders as ODD and Conduct Disorder (CD; Keenan & Wakschlag, 2002). Application of the diagnostic criteria also raises questions regarding whether such requirements as cross-situational impairment and duration of symptoms are relevant and appropriate for preschoolers (Byrne et al., 2000; Lahey et al., 2004). Such concerns highlight the current lack of sufficient information about the feasibility or validity of applying standard diagnostic procedures to the preschool population (Egger & Angold, 2006; Gadow, Sprafkin & Nolan, 2001).

This study examined the ways in which dimensional measures of behavior differentiated among preschool-aged children who met criteria for ADHD, with or without comorbid ODD. Various measures of behavior were examined, including parent and preschool teacher ratings on measures of behavior, attention, and social skills, as well as direct behavioral observation. Further analysis examined the extent to which these empirically-defined subgroups corresponded with diagnostic classifications. Specific research questions are described below, along with the hypothesized outcomes.

1. In what way(s) do dimensional measures of behavior, including the Child Behavior

Checklist (Achenbach, 1991a), Teacher Report Form (Achenbach & Edelbrock, 1986), Social Skills Rating System (Gresham & Elliott, 1990), Conners' Rating Scales (Conners, 1997), and direct observation differentiate distinct subgroups within a larger group of preschoolers who meet criteria for ADHD with or without comorbid ODD?

   a. It is hypothesized that cluster analysis will identify a greater number of homogeneous subgroups among the participants than would directly correspond with ODD and/or the various subtypes of ADHD as defined in the DSM-IV.

2. What salient characteristics will differentiate among the identified subgroups?

   a. In direct contrast with a categorical (symptom present/absent) approach to diagnosis, it is hypothesized that the severity of behavioral and/or emotional characteristics will contribute to the distinction among subgroups. Severity will be determined by examination of the mean score profile of each cluster on the scales/ measures included in analysis.

   b. Based on prior evidence of inconsistency in parent and teacher ratings (Achenbach et al., 2008; Drabick, Gadow, Carlson & Bromet, 2004; Jensen et al., 1993; Serna et al., 2002), it is expected that parent and teacher ratings characterizing the various clusters will differ.

3. Will classifications based on cluster analysis be consistent with classifications based on DSM-IV diagnostic categories?

   a. Consistent with previous findings of a relationship between rating scale measures and DSM diagnoses (Biederman et al., 1993; Chen et al., 1994; Edelbrock & Costello, 1988; Frankel, Hanna, Cantwell, Shekim & Ornitz, 1992; Hudziak et al., 2004), it is expected that clusters will share some similarities with current

diagnostic criteria (e.g., patterns of elevation on scales and measures corresponding with symptom characteristics in the DSM-IV).

b. It is expected that participants with the same diagnosis(es) will tend to be grouped within the same cluster (Frankel et al., 1992).

The current study will contribute to the existing knowledge base in several ways. Knowledge about behavioral disorders at the preschool level will be enhanced through analysis of a sample comprised exclusively of preschool-aged children. The relationship between psychiatric disorders as defined in the DSM-IV and dimensional measures of behavior will be further clarified through inclusion of a variety of assessment methods and sources of information. In contrast with previous studies assessing the relationship between DSM-based and dimensional classifications, the current investigation will analyze the degree of similarity between the characteristics of empirically-derived subgroups and the symptom patterns used to define various disorders. Reliance on an empirical approach to classification may illustrate unanticipated differences among children, with potential implications for their current functioning and later outcomes.

Chapter II: Literature Review

Neurodevelopmental disorders (NDs) and disruptive behavior disorders (DBDs) are widely recognized as prevalent among school-aged children; however, research has indicated that problem behaviors emerge at an early age. A significant number of preschoolers exhibit symptoms characteristic of NDs and/or DBDs (Gimpel & Kuhn, 2000; Lavigne et al., 1996). Those most prevalent among young children are Attention Deficit Hyperactivity Disorder (ADHD) and Oppositional Defiant Disorder (ODD). Prevalence rates of ADHD and ODD within the preschool population have been found to be highly similar to the rates evident among school-aged children (Lavigne et al., 1996).

Temporal Stability of Early-Onset Behavior Problems

The importance of early identification and intervention is underscored by findings that early-onset behavior problems tend to remain stable over time. Campbell and colleagues conducted a series of longitudinal studies to examine the temporal stability of early problem behaviors and to identify predictors of later outcomes. The target sample consisted of three-year-olds recruited via parent referral for concerns about five symptoms: "a high level of motor activity, short attention span, difficulty playing alone, tantrums, and defiance" (Campbell et al., 1984, p. 244). Data were also gathered on a control comparison group of same-aged children.

At 11-month follow-up, the mothers of "problem children" were more likely to report behavior problems than the mothers of children in the control group (Campbell et al., 1984). Maternal interviews indicated higher levels of restlessness, impulsivity, inattention, tantrums, problems in preschool, and problems getting along with peers and siblings among the target children. These problems were more likely to be rated as moderate to severe by mothers of the referred children, whereas children in the control group were more likely to be rated as

exhibiting no problems or mild/age-appropriate problems. Maternal questionnaires also reflected higher levels of aggressive-defiant and hyperactive-distractible behavior among referred children.

Parent reports of problem behaviors were supported by direct observation of the children's behavior during free play and structured tasks. During free play, the children referred for problem behaviors exhibited a higher activity level, more frequent shifts between activities, briefer engagement with toys, and more time engaged with non-toy objects ("primarily forbidden objects such as locked cabinets, doors to the video-camera, and one-way screen;" Campbell et al., 1984, p. 245) than control comparison. During structured activities, children referred for problem behaviors exhibited more out-of-seat behavior and a higher number of impulsive responses to a task requiring response delay than children in the control group.

A follow-up study conducted three years after the initial study indicated that fully half of the children initially referred for behavioral problems continued to exhibit elevated problem behaviors at age six (Campbell, Ewing, Breaux & Szumowski, 1986b). Children were identified as having "continuing" or "persistent problems" if they met one of three criteria: (a) meeting diagnostic criteria for ADHD, (b) receiving a maternal rating of aggression within the clinically significant range, or (c) being rated as having moderate or severe problems on a maternal structured interview. Children who had initially been referred for behavior problems, but who did not meet any of the above criteria at follow-up were classified as "improved."

The results reflected a high level of continuity between early behavior problems (at age three) and symptom manifestation at home and in school three years later, at age six (Campbell et al., 1986b). At follow-up, referred children differed from control comparisons on several parent-report, teacher-report, and direct observation measures. Maternal ratings reflected

significantly higher levels of hyperactivity, aggression, social withdrawal, and peer problems, as well as lower social competence among referred children compared with children in the control group. Teacher ratings indicated that children referred at age three exhibited significantly more inattention, overactivity, and aggression, as well as poorer adaptive functioning in school at age six than their non-referred peers. Direct observation of children during structured classroom activities approached significance, with referred children exhibiting somewhat more aggressive-disruptive and hyperactive-distractible behaviors than control comparisons, although the results were not statistically significant ($p < 0.06$ and $p < 0.08$, respectively).

According to maternal interview and ratings, 50% of children who had been identified as "hard-to-manage" at age three were determined to have persistent (rather than "improved") problems at age six (Campbell et al., 1986b). Compared with children whose problem behaviors persisted over time, those whose behavior problems had improved by age six were rated by mothers and teachers as having lower levels of externalizing behavior. They also exhibited less inattentive and uncooperative behavior during direct observation. Children in the "improved" group did not differ significantly from control comparisons at age six, suggesting that their behavior had "normalized" to a great extent.

The results of these longitudinal studies by Campbell and colleagues indicate that, for a substantial number of children, early-onset problem behaviors remain stable over time. Parent ratings indicated that early problem behaviors continued to be perceived as moderate to severe at one- and three-year follow-up assessments. Teacher ratings and direct observations in the classroom setting confirmed that children referred as three-year-olds exhibited more problem behavior at age six than their typical peers.

Relationship between Early-Onset Behavior Problems and Subsequent Diagnostic Status

In addition to findings supporting the long-term stability of early-onset problem behaviors, longitudinal studies have also produced findings that children who exhibit behavior problems at a young age have a high likelihood of subsequently meeting diagnostic criteria for a neurodevelopmental disorder (ND). In the longitudinal study by Campbell and colleagues (1986b) described above, the majority of children whose problem behaviors persisted from age three (63%) subsequently met diagnostic criteria for ADHD (per structured maternal interview) at age six. Several measures administered at age six reflected significant differences between children who met criteria for ADHD compared with children in the "improved" and control groups.

After analyzing group differences evident at age six follow-up, Campbell and colleagues (1986b) examined whether baseline measures (conducted at age three) differentiated among subgroups that emerged subsequently, at age six. Several direct observation measures (including behavior during free play, out-of-seat and off-task behavior during a structured task, and impulsive responses during a response delay task) differentiated referred children from control comparisons, but did not differentiate children whose problems later improved from those who subsequently met DSM-III criteria for ADHD. However, maternal ratings at baseline did distinguish among the three subgroups that emerged three years later. Children who met criteria for ADHD at age six were rated as significantly more hyperactive and impulsive at age three than children in the improved group, who, in turn, were rated as more hyperactive than control comparisons. Children in the ADHD group were also rated by their mothers as more aggressive and defiant at baseline than control comparisons. Maternal ratings of aggression and defiance for children in the improved group were lower than ratings for the ADHD group and higher than the

control group, but the differences were not statistically significant.

Campbell and Ewing (1990) continued to follow the same cohort of children originally recruited at age three based on parent concerns regarding limited attention span, high activity level, tantrums, defiant behavior, and difficulty playing alone. Data were gathered six years later (at age nine) from mothers and teachers, providing information about the children's functioning across settings. Parent measures included a structured maternal interview, the parent version of the NIMH Diagnostic Interview Schedule for Children (DISC-P; Costello, Edelbrock, Kalas, Kessler & Klaric, 1982), the CBCL, the SNAP Questionnaire (evaluating for symptoms of ADHD as well as problems with peers), and the Life Experiences Survey (Sarason, Johnson & Siegel, 1978). Teacher-report measures included the Teacher Report Form (TRF, Achenbach & Edelbrock, 1986) and the SNAP Questionnaire.

Consistent with previous findings, the assessments completed by Campbell and Ewing (1990) revealed continued behavior problems through age nine, with a substantial number of children meeting diagnostic criteria for neurodevelopmental disorders (NDs) and/or disruptive behavior disorders (DBDs). Among children whose early behavior problems persisted through age six, 67% met diagnostic criteria for an ND and/or DBD at age nine (per the DISC-P). Subsequent analysis revealed convergence between CBCL ratings and DSM-III diagnostic criteria. The CBCL Externalizing Scale had moderate to strong correlations with the overall number of externalizing symptoms ($r = 0.72$) and with the ADDH scale ($r = 0.68$) of the DISC-P. Correlations between the CBCL Externalizing Scale and other DISC-P diagnostic scales were also statistically significant, though not as strong (Conduct Disorder: $r = 0.42$, Oppositional Defiant Disorder: $r = 0.55$). Conversely, subscales of the CBCL were also found to have a significant correlation with the overall, externalizing symptom count of the DISC-P

(Hyperactive: $r = 0.60$, Aggressive: $r = 0.66$, Delinquent: $r = 0.60$), reflecting a correspondence between rating scale scores and reported ND/DBD symptoms.

Continuity in symptoms over time was associated with increased functional impairment, as assessed by parent and teacher report as well as use of special services (Campbell & Ewing, 1990). In comparison with children in the "improved" and control groups, children whose problem behaviors were identified as "persistent" from age three to age six were rated (by parents) at age nine as exhibiting: more symptoms of ADHD, more peer problems, higher levels of aggression and delinquency, more hyperactivity, and less social competence. Consistent with findings reported by Campbell and colleagues (1986b) at the three-year follow-up (age six), Campbell and Ewing (1990) found that at the six-year follow-up (age nine), children with persistent problems were more likely to receive remedial academic services or special education services and were more likely to access services in the community (e.g., psychiatric or psychological assessment, consultation with a pediatrician) than children in the control or "improved" groups. Teacher ratings indicated that at age nine, children with persistent problems exhibited more ADHD symptoms, more externalizing problems (e.g., nervous-overactive behavior), and more peer problems than control comparisons; however, teacher ratings did not significantly differentiate children in the persistent problem group from those whose early problem behaviors had improved.

Analyses were also conducted to examine which of the earlier measures predicted subsequent ND/DBD symptoms (Campbell and Ewing, 1990). Various measures at age three, including parent-report measures as well as children's behavior during direct observation, were predictive of ND/DBD symptoms six years later. Specifically, mother-reported ADHD symptoms at age nine (per the DISC-P) were predicted by various maternal reports (difficult infant temperament,

composite hyperactivity rating, and SNAP rating) and by composite "free play" score at age three (Campbell & Ewing, 1990). Similarly, mother-reported symptoms of ODD and CD at age nine were predicted by maternal reports (difficult infant temperament and Behar aggression scale) as well as direct observation (free play composite score) at age three. Teacher-reported externalizing behaviors at age nine (via the TRF) were predicted by children's observed behavior at age three (inattention and activity level during free play) and by teacher ratings of externalizing and internalizing behavior at age six (completed by different teachers via the TRF).

In a subsequent follow-up study focusing on symptom manifestation and diagnostic status, Pierce and colleagues (1999) gathered data on the same cohort of children at age 13 (Cohort 1). Data were also gathered on a cohort of nine-year-old boys who had also been identified as "hard-to-manage" during preschool by parents and/or preschool teachers and who were subsequently assessed at age six (Cohort 2). Parent-report measures for Cohort 1 included a structured, diagnostic interview (the Child Assessment Schedule, Parent Form (P-CAS); Hodges, Saunders, Kashani, Hamlett & Thompson, 1990) as well as the CBCL rating scale (both completed by mothers). Children completed the Youth Self-Report rating scale (Achenbach & Edelbrock, 1987). Measures for Cohort 2 included the SNAP questionnaire (completed by preschool teachers at intake and by parents at assessment ages four, five, six, and nine), the CBCL (completed by mothers at each follow-up and by fathers at the age nine assessment), and the P-CAS (administered to mothers at the age nine assessment).

Consistent with previous findings, children identified as "hard-to-manage" at age three exhibited more externalizing problems at age 13 (Cohort 1) than control comparisons (Pierce et al., 1999). Maternal CBCL ratings indicated higher levels of aggression, delinquency, and overall externalizing problems among referred children at age 13, compared with those in the

control group. “Hard-to-manage” boys were also rated as exhibiting higher levels of hyperactivity than their comparison peers, although this difference was not significant among girls. No significant differences were found between groups with regard to parent-rated social competence or on the self-report measure, although the authors reported a trend toward higher levels of self-reported aggression and delinquency and lower levels of social competence among referred children.

Children identified as “hard to manage” at age three (Cohort 1) were also more likely to meet diagnostic criteria at age 13 than their non-referred peers (Pierce et al., 1999). Fifty-three percent (53%) of referred children met diagnostic criteria for ADHD at age 13, compared with only 2% of control comparisons. Similarly, “hard to manage” children were more likely to meet diagnostic criteria for ODD (26%) or CD (15%) at age 13 than were non-referred controls (8% and 0%, respectively). Children whose symptoms were stable from an early age had a higher rate of oppositional/defiant behavior and a greater likelihood of subsequently meeting diagnostic criteria for ADHD and for comorbid diagnoses than children whose early behavior problems showed improvement over time. Among children whose behavior problems persisted from age three through ages six and nine, 94% met diagnostic criteria at age 13, compared with fewer than 25% of children whose early behavior problems improved over time.

A similar pattern was evident among participants in Cohort 2. Boys identified as “hard-to-manage” as preschoolers were rated by both parents as exhibiting more social problems, higher levels of aggression, more attention problems, and more overall externalizing problems at age nine than boys in the non-referred comparison group (Pierce et al., 1999). “Hard-to-manage” children in Cohort 2 were also more likely to meet diagnostic criteria for an ND and/or DBD than children in the comparison group (34% versus 13%). A greater proportion of children

referred at age three met diagnostic criteria for ODD/CD at age nine (28%) than their non-referred peers (8%). Similarly, a greater number of boys in the "hard-to-manage" group met diagnostic criteria for ADHD (19% versus 5% of control comparisons), although the difference was not statistically significant. Examination of the prevalence of comorbid disorders also revealed a higher rate of combined ADHD and ODD/CD among "hard-to-manage" boys (n = 11) than among those in the comparison group (n = 1), although these results also failed to reach statistical significance.

Consistent with the findings for children in Cohort 1, symptom chronicity over time was found to be a significant factor differentiating boys in Cohort 2 who went on to meet criteria for an externalizing disorder from those who did not (Pierce et al., 1999). Children whose behavior problems persisted through age six were 4.5 times more likely to have an externalizing disorder at age nine than those whose symptoms had improved by age six. At baseline and at each follow-up assessment, maternal ratings (CBCL Externalizing and SNAP scores) were significantly higher for children with persistent problems than for children whose behaviors improved, or control comparisons. Problem persistence was also associated with receiving an elevated CBCL Internalizing score at some point during follow-up assessments. These results indicate that symptom severity is an important factor in predicting symptom continuity and later outcomes (including diagnostic status).

Overall, the series of studies conducted by Campbell and colleagues and by Pierce and colleagues indicated that a significant proportion of children who were identified as "hard-to-manage" at age three continued to exhibit elevated behavior problems throughout a succession of follow-up assessments, up to 10 years later. Significant differences between referred children and control comparisons were reflected in various measures, including parent report, teacher report,

and direct observation. Referred children were more likely subsequently to meet criteria for an externalizing disorder, demonstrating continuity in behavior problems over time and the relationship between early behavior problems and subsequent diagnostic status. Examination of baseline differences between subgroups indicated that some early behavior measures (primarily parent report) distinguished children whose early onset problems later improved from those whose problems persisted over time.

Lahey and colleagues (1998) also conducted a series of studies examining the long-term trajectory of symptom manifestation. Children ages three through seven were recruited in Chicago and Pittsburgh (the majority were between the ages of four and six). Participants were recruited through flyers and from university child psychiatry clinics, where they presented with symptoms of inattention and hyperactivity. Control comparisons were recruited from the same or similar schools as those attended by the children exhibiting symptoms of ADHD.

An early study focused on the validity of the DSM-IV ADHD diagnosis for younger children (Lahey et al., 1998). Children were classified as meeting criteria for ADHD if symptoms were reported by the parent (via structured interview using the Diagnostic Interview Schedule for Children, Version 2.3) or by the teacher (via the DSM-IV version of the DBD Checklist; Pelham, Gnagy, Greenslade & Milich, 1992). Potential impairment in functioning was assessed through parent- and teacher-report as well as child self-report.

The findings indicated greater impairment (per teacher rating) in social skills and in adaptive functioning among children who met criteria for any subtype of ADHD compared with children in the control group (matched for age, gender, and ethnicity; Lahey et al., 1998). After controlling for several potentially confounding variables (symptoms of ODD and CD, parent-reported internalizing symptoms, age, gender, family income), teacher ratings reflected greater

social impairment among children who met diagnostic criteria for each subtype of ADHD, including lower scores on scales assessing prosocial skills, cooperation, and assertion. Children who met diagnostic criteria were also rated by their teachers as being ignored by more classmates and being liked by fewer classmates than children in the control group.

Similar findings were evident in self-report and parent rating measures (Lahey et al., 1998). Self-reports by children who met criteria for any subtype of ADHD reflected significantly more difficulties with friendships than self-reports completed by comparison children. Parent and interviewer ratings also indicated that control comparisons exhibited higher levels of adaptive functioning than children in any of the three ADHD subgroups. Children who met criteria for any subtype of ADHD were also more likely than control comparisons to have received special education services. Overall, the results indicated that young children who met criteria for any subtype of ADHD exhibited impairments in various aspects of functioning compared with typical peers.

In a follow-up study, Lahey and colleagues (2004) examined the predictive validity of an early ADHD diagnosis (made at the time of first assessment) across three subsequent annual assessments. Upon initial assessment, a substantial number of participants who met symptom criteria for an ADHD diagnosis exhibited impairment in only one setting (22.5%), failing to meet the requirement for cross-situational impairment. These participants were classified into a "situational ADHD" group to allow for analysis of the validity of including cross-situational impairment among the diagnostic criteria for young children. Among children who met full diagnostic criteria ("full ADHD"), 76% met criteria for the Combined Type of ADHD, in contrast with only 29.4% of those within the "situational ADHD" group. The authors noted that for this reason, comparison between the two groups also amounted to a comparison between

ADHD subtypes.

Follow-up assessments included parent-, teacher-, and child self-report measures (Lahey et al., 2004). The results provided general support for the validity of the DSM-IV diagnostic criteria for ADHD among young children, but raised questions regarding the appropriateness of the requirement for cross-situational impairment. Children who initially met full criteria for ADHD (at ages three to seven) were likely to continue to meet full diagnostic criteria across all subsequent assessments. Children in the "situational ADHD" group were also likely to meet full diagnostic criteria at least twice during subsequent assessments, although less frequently than children who initially met full criteria. Children in both ADHD groups exhibited greater impairments in adaptive and social functioning than control comparisons, across parent-, teacher- and self-report ratings. Although the impairments were more pronounced for children in the "full ADHD" group, children in the "situational ADHD" group also exhibited greater impairments in functioning and a higher likelihood of subsequently meeting full diagnostic criteria, raising questions regarding the appropriateness of requiring impairment across settings in order to diagnose ADHD among very young children.

In a subsequent study examining the temporal stability of ADHD diagnoses among the same participants, Lahey and colleagues (2005) found that although the log odds of meeting criteria for ADHD declined over time, a substantial number of children continued to meet diagnostic criteria. Children who met criteria for the Combined Type of ADHD at initial assessment were most likely to continue meeting diagnostic criteria for any ADHD subtype over time, with no significant difference between the Hyperactive-Impulsive or Inattentive subtypes. The authors noted that the overall stability in diagnosis was obscured by variability in symptom manifestation over time.

When diagnostic stability was defined as meeting criteria for any ADHD subtype during at least one of the last two follow-up assessments, ADHD diagnosis was found to be generally stable, with no significant differences among the three subtypes (Lahey et al., 2005). However, the majority of children in each diagnostic subgroup met criteria for a *different* ADHD subtype during at least one follow-up assessment. Over time, the prevalence of the Combined Type (CT) and the Hyperactive-Impulsive Type (HT) of ADHD decreased, while the prevalence of the Inattentive Type (IT) increased. Further examination of children's mobility between subtypes revealed that of the children initially in the Hyperactive-Impulsive group who continued to meet diagnostic criteria in either of the last two assessments, very few remained in the HT group; the majority met criteria for the Combined Type of ADHD. Of the 64 children who initially met criteria for the Combined Type of ADHD and who continued to meet criteria for ADHD during one of the last two assessments (year seven or eight), the majority (75%) again met criteria for the Combined Type during one of those years. Among the same group of children, a slight majority (51%) met criteria for the Inattentive Type during year seven or eight. Most of the children initially classified in the IT group (six out of eight, or 75%) remained in this group during at least one of the last two assessments. The other two children met criteria for CT or HT in one of the final assessment years.

Analysis of symptom manifestation revealed several changes over time (Lahey et al., 2005). The average number of symptoms of hyperactivity-impulsivity decreased over time, but symptom patterns across subgroups were theoretically consistent. Children in the HT and CT subgroups exhibited similar mean levels of hyperactive-impulsive symptoms, both displaying higher levels of these symptoms than children in the IT subgroup. There was no significant overall reduction in symptoms of inattention over time. Consistent with diagnostic theory,

children in the HT subgroup exhibited fewer symptoms of inattention across the follow-up assessments than children in the CT or IT subgroups, while the latter two groups did not differ with regard to symptoms of inattention. However, there was a significant interaction effect. Over time, symptoms of inattention slightly increased among children in the HT group, but decreased among children in IT and CT groups.

Overall, the longitudinal studies conducted by Campbell, Lahey, and colleagues demonstrate the temporal persistence of early behavior problems and their strong relationship with subsequent diagnoses. Developmental changes and/or environmental demands may interact to affect symptom manifestation over time; however, a substantial number of children who exhibited symptoms of ADHD at an early age continued to meet diagnostic criteria over time. For children with comorbid disorders, the risk for symptom stability and severity is even greater.

Impact of Comorbidity

The risk for symptom stability and severity is particularly elevated among children with comorbid disorders (Speltz et al., 1999; Tandon, Si & Luby, 2011). Research of school-aged children indicated that those who met criteria for both ADHD and ODD exhibited more severe ADHD symptoms (Abikoff et al., 2002; Biederman et al., 2006; Newcorn et al., 2001), more oppositional behavior (Biederman et al., 2006), and higher levels of aggression and "interference" behaviors (e.g., interrupting the teacher; Abikoff et al., 2002) than children with ADHD alone. The relationship between comorbidity and increased breadth and severity of symptoms has also been documented among participants of a broader age range.

In an early study of children between the ages of 5 and 12, Kuhne and colleagues (1997) found that those with ADHD and a comorbid DBD (either ODD or CD) exhibited a greater number of ADHD symptoms than children with a sole diagnosis of ADHD, as measured by

diagnostic parent interview. Children with comorbid ND/DBDs were also described by parents and teachers as significantly more likely to be irritable or to fight with friends frequently than were children with ADHD alone. Parents further reported that children with ADHD and comorbid ODD or CD engaged in mutual activities with other children less often than children with ADHD only.

Similarly, in a meta-analysis including nearly 25,000 participants between the ages of 4 and 20, Waschbusch (2002) found evidence of greater symptom severity among individuals with comorbid disorders. The purpose of the review was to examine the functional impact of hyperactive-impulsive-attention problems (HIA), conduct problems (CP), and/or a combination of both (HIA-CP), as well as the relationship(s) among these three patterns of symptomology. The nature of the symptomology manifested by individuals with and without comorbid diagnoses was analyzed after averaging across parent and teacher reports (via behavior ratings or structured interview). The results revealed that participants with symptoms of both ADHD and conduct problems (HIA-CP) exhibited more severe hyperactivity, impulsivity, conduct problems, and aggression than those with symptoms in only one category of behavior (or control comparisons).

Comparisons between individuals exhibiting only one category of symptoms were theoretically consistent with diagnostic criteria. Those in the HIA group were described by parents and teachers as demonstrating higher rates of hyperactivity and impulsivity than children with conduct problems only (CP), but fewer conduct problems (Waschbusch, 2002). On ratings of inattention, individuals in the HIA group were rated the highest, followed by the HIA-CP group, and finally the CP group, all with higher levels of inattention than control comparisons.

Consistent with the results reported by Kuhne and colleagues (1997), the meta-analysis conducted by Washbusch (2002) also suggested a relationship between comorbidity and various

types of social interaction difficulties. Among the studies reviewed, three used performance measures to assess for various aspects of antisocial behavior: reactive aggression (aggression that occurs in response to provocation), instrumental aggression (which involves achieving a reward or advantage through an act of aggression), and hostile aggression (which involves inflicting harm on another for no apparent reward or advantage). The results revealed more significant problems among individuals with comorbid ND/DBDs than those exhibiting a single disorder.

Individuals with symptoms of ADHD as well as conduct problems (HIA-CP group) exhibited higher levels of reactive aggression in response to minimal provocation and longer maintenance of aggression over time than children who exhibited symptoms in only one category (HIA or CP; Waschbusch, 2002). A study comparing instrumental and hostile aggression found that children with conduct problems only (CP) and those with both conduct problems and ADHD symptoms (HIA-CP) exhibited significantly higher rates of instrumental aggression than control comparisons, but only the comorbid group differed significantly from controls with regard to hostile aggression. Children with symptoms of ADHD and conduct problems had higher levels of both instrumental and hostile aggression than children in the conduct problems-only group, but the differences were not statistically significant.

Eleven of the studies reviewed (Washbusch, 2002) examined social functioning as assessed by negative peer nominations (seven studies), social preference (two studies), and peer ratings of liking (two studies). Overall analysis of effect sizes across the studies indicated that children exhibiting symptoms of both ADHD and a comorbid DBD had significantly more difficulties with their peers than any other group. Children exhibiting symptoms of ADHD-only had significantly more peer difficulties than children in the DBD-only group.

Although the results of Waschbusch's (2002) meta-analysis are thought-provoking, it is

important to note that methodological differences across the reviewed studies affected the results. Both the nature of the samples and the method(s) used to group participants into different diagnostic categories impacted the resulting differences between participant subgroups. Compared with studies using different grouping criteria, those that used DSM criteria to classify participants found significantly more peer difficulties among children in the comorbid group (HIA-CP) than children in the ADHD-only group. In addition, groupings based on DSM criteria showed a less substantial difference in level of peer difficulty in ADHD-only children compared with the control group.

Examination of differences in results across sample types revealed that in studies using referred samples, the differences between children in the HIA-CP groups and those in the control groups were larger than the differences between these groups in non-referred samples. Similarly, children in the HIA-CP group were significantly different from children in the HIA-only group among referred samples, but not in non-referred samples. The magnitude of the differences among symptom-based subgroups was affected not only by sample source and basis for group assignment, but also by the measurement(s) used. Use of the Aggression subscale of the CBCL as a measure of conduct problems resulted in significantly larger effect sizes than did use of the Delinquency subscale (Waschbusch, 2002).

Despite the impact of variations in methodology, collectively, the research indicates that children with comorbid ND/DBDs exhibit greater functional impairment than children with only one ND or DBD diagnosis, including problems across a wider range of domains, as well as elevation in the number and severity of symptoms. However, these results reflect findings for school-aged children and children across wide age ranges. Developmental differences and the impact of the environment on symptom manifestation compel specific examination of the

characteristics exhibited by preschoolers with ND/DBD diagnoses before conclusions can be drawn regarding the applicability of these findings to a younger population.

Symptom Manifestation in the Preschool Population

In a study conducted exclusively with preschool-aged participants, Cunningham & Boyle (2002) evaluated the social and behavioral characteristics of a community sample of four-year-olds enrolled in "universal junior kindergartens" (Cunningham & Boyle, 2002, p. 557). Parent and teacher responses to a rating scale based on the DSM-III-R were used to group the children according to whether they were at risk for: ADHD only, ODD only, or comorbid ADHD and ODD (ADHD+ODD). Children with scores below the 1.5 standard deviation cut-off were included in the "Normal" group.

The majority of statistically significant findings pertained to differences between children in one of the at-risk subgroups compared with children not at risk for an ND or DBD (Cunningham & Boyle, 2002). Children in every at-risk group (ADHD-only, ODD-only, or ADHD+ODD) were rated by their parents as having more severe symptoms than children in the "Normal" group. Children classified as being at risk for ADHD-only were rated by parents and teachers as having more attention problems and more social problems than children in the "Normal" group. In addition, children at risk for ADHD-only were rated by their parents as having more internalizing problems and as being problematic in more situations; they were rated by teachers as more aggressive than children in the "Normal" group.

Differences between children at risk for ODD-only and those in the "Normal" group were largely limited to parent ratings and direct observation of parent-child interactions (Cunningham & Boyle, 2002). Parents rated children in the ODD-only group as more aggressive and as having more internalizing problems than children in the "Normal" group. Parents of children in the

ODD-only group also reported poorer family functioning and a lower sense of competence than parents of children who did not exhibit symptoms of ODD. During observation, children at risk for ODD-only complied with fewer parent directives and engaged in lower levels of compliant behavior than children in the "Normal" group.

Differences between children identified as being at risk for comorbid ADHD+ODD and those in other diagnostic subgroups varied across parent and teacher ratings (Cunningham & Boyle, 2002). Parent-based measures reflected significant differences primarily between children in the ADHD+ODD group and those in the ADHD-only group. According to parent ratings, children categorized as at risk for both ADHD and ODD were rated as being problematic in more situations than children in the ADHD-only subgroup. Parent ratings trended toward higher levels of internalizing problems and lower parent self-reported sense of competence in the ADHD+ODD group compared with the ADHD-only group, but these results were not statistically significant.

Teacher-based measures primarily reflected differences between children in the ADHD+ODD group and those in the ODD-only group (Cunningham & Boyle, 2002). Teachers rated children at risk for both ADHD and ODD as having more social problems and not working as hard as children at risk for ODD-only. One significant finding consistent across both parent and teacher ratings was that children in the ADHD+ODD group were rated as having more attention problems than children in the ODD-only group.

Overall, the differences among preschoolers at risk for various ND/DBDs generally corresponded with diagnostic criteria, although the differences varied across respondent (Cunningham & Boyle, 2002). Children at risk for ADHD-only were rated by teachers as having more attention problems and more social problems than children in the ODD-only group.

Children in the ODD-only group were rated by parents as more aggressive than children in the ADHD-only group. During parent-child interactions, children at risk for ODD-only engaged in lower levels of compliant behavior than children in the ADHD+ODD group. Parent and teacher ratings of attention were not significantly different for the ODD-only group compared with children who did not exhibit symptoms of ODD and those in the "Normal" group.

Gadow and Nolan (2002) also found that parent and teacher ratings of the symptoms exhibited by preschoolers in different ND/DBD-based subgroups yielded somewhat divergent results. Parent and teacher ratings on a scale assessing the presence and severity of symptoms of various disorders (as defined in the DSM-IV) were analyzed separately, examining differences between subgroups of children ages three to six classified as exhibiting symptoms of ADHD-only, ODD-only, ADHD+ODD, or symptoms of neither disorder (Comparison group). The analysis was completed once with a community sample and once with a clinic-referred sample. Within the community sample, analysis of parent-defined and parent-rated subgroups indicated that children classified as exhibiting symptoms of both ADHD+ODD were rated as exhibiting a higher overall number of symptoms, an increased level of symptom severity overall, as well as higher scores (reflecting more significant symptom severity) in several individual symptom domains (e.g., Generalized Anxiety Disorder, Dysthymic Disorder, Major Depressive Disorder, peer conflict), compared with children in the ADHD-only, ODD-only, and Comparison groups. Parent ratings tended not to distinguish between children in the ADHD-only and ODD-only subgroups. The only exceptions to this pattern pertained to symptoms of Conduct Disorder and the total number of symptoms reported overall; in both cases, children with ODD-only were rated as having more problems than children in the ADHD-only group. Across nearly all analyses of parent ratings, children in the Comparison group were rated as exhibiting the least

severe symptoms.

Teacher ratings of the community sample were somewhat more sensitive to between-group differences (Gadow & Nolan, 2002). Consistent with the results based on parent ratings, children within the teacher-defined ADHD+ODD subgroup were rated by their teachers as exhibiting a higher overall number of symptoms and a higher level of symptom severity overall, in comparison with each other subgroup. In contrast with parent ratings, teacher ratings were also more elevated for children in the ADHD+ODD subgroup with regard to symptoms of Autistic Disorder and Asperger's Disorder. In numerous areas, teacher ratings of children in the ADHD+ODD subgroup and ODD-only subgroup were similar, both rated as exhibiting more severe problems than children in the ADHD-only or Comparison groups (e.g., Conduct Disorder, Generalized Anxiety Disorder, Major Depressive Disorder, Dysthymic Disorder, Developmental Deficits, and peer conflict).

For the clinic-referred sample, parent ratings reflected similarities between the ADHD+ODD and ODD-only subgroups, which were rated as exhibiting a higher overall number of symptoms, an increased level of symptom severity overall, as well as higher scores (reflecting more significant symptom severity) in several individual symptom domains (e.g., Conduct Disorder, Generalized Anxiety Disorder, Major Depressive Disorder, and Dysthymic Disorder), compared with children in the ADHD-only and Comparison groups (Gadow & Nolan, 2002). The domains that distinguished between children in the ADHD+ODD subgroup and the ODD-only subgroup pertained to symptoms of Autistic Disorder and Asperger's Disorder. Preschoolers in the ADHD-only group were not significantly different from those in the Comparison group in most areas assessed, per parent ratings.

Teacher-defined and –rated subgroups of clinic-referred children also indicated greater

impairment among children exhibiting symptoms of both ADHD+ODD compared with children in the ADHD-only and Comparison groups, with regard to total number of symptoms and overall symptom severity (Gadow & Nolan, 2002). Children in the ADHD+ODD subgroup were also rated by teachers as exhibiting more severe symptoms in the areas of Conduct Disorder, Generalized Anxiety Disorder, Major Depressive Disorder, and Dysthymic Disorder. Teacher ratings did not significantly distinguish between children with ODD-only and those with ADHD-only, with the exception of symptoms Conduct Disorder, which were more elevated among the former subgroup.

Overall, the findings indicate that children exhibiting symptoms of both ADHD and ODD tended to be rated by both parents and teachers as more impaired than children with symptoms of ADHD-only or control comparisons (Gadow & Nolan, 2002). Other between-group differences were less consistent, varying according to respondent and sample source (community versus clinic-referred). Teacher ratings of a community sample of preschoolers tended to follow the same pattern as parent ratings of a clinic-referred sample, reflecting similarities in the symptoms exhibited by children with ADHD+ODD and those with ODD-only, both rated as having more difficulties than children with ADHD-only or those in the comparison group.

A more recent study of preschoolers by Healey and colleagues (2011) examined the relationship between ADHD symptom severity and negative emotionality, and the extent to which the relationship between these factors might be mediated by cognitive functioning. Children between the ages of three and four were recruited through local preschools. The participants were classified either as typically developing or as hyperactive/inattentive, based on screening measures completed by parents and teachers. Symptoms consistent with ADHD were rated by parents; teacher ratings were used to measure negative emotionality (described as “a

tendency to become emotionally distressed," Healey, Marks & Halperin, 2011, p. 504). The Developmental Neuropsychological Assessment (NEPSY; Korkman, Kirk & Kemp, 1998) was administered to gather information about the participants' cognitive/neurological functioning. The children's performance was evaluated along two factors suggested by principal components analysis: *Verbal and Executive Functions* (VEF) and *Perceptual-Motor and Executive Functions* (PMEF).

Significant correlations were found among all of the variables (Healey et al., 2011). Higher levels of negative emotionality (per teacher ratings) were associated with heightened ADHD symptom severity (per parent ratings) and with lower levels of functioning on the two NEPSY factors: VEF and PMEF. Heightened severity of ADHD symptoms was also independently associated with lower performance on the VEF and PMEF measures.

The relationships among the variables were further examined through hierarchical linear regression analysis (Healey et al., 2011). The results indicated that for children with higher levels of negative emotionality, functioning on the VEF and PMEF did not have a significant relationship with ADHD symptomology. However, when negative emotionality was low, weaker performance on the VEF and on the PMEF was (each independently) associated with increased severity of ADHD symptoms. In addition, the relationship between negative emotionality and ADHD symptoms was stronger among children with higher VEF and PMEF scores. Follow-up analysis did not reveal a difference in the nature of these relationships when symptoms of inattention and symptoms of hyperactivity/impulsivity were considered separately.

These results suggest that among preschoolers, both (teacher-rated) negative emotionality and cognitive functioning are associated with the severity of parent-rated ADHD symptoms (Healey et al., 2011). The directions of the relationships were as might be predicted, with higher

levels of negative emotionality and lower levels of cognitive functioning (across two domains) associated with increased symptom severity. The moderating effect of negative emotionality suggests that when negative emotionality is low, higher cognitive functioning is associated with lower ADHD symptom severity. However, this apparent protective factor was not evident among preschoolers with high levels of negative emotionality. Preschoolers rated by teachers as exhibiting high levels of negative emotionality were rated by their parents as having the most severe symptoms of ADHD, and their level of cognitive functioning was not significantly related to ADHD symptom severity.

Together, these studies suggest that among preschoolers, several factors may impact symptom manifestation. The presence of comorbid disorders and negative emotionality appear associated with increased symptom severity. Cognitive functioning may also impact the severity of ADHD symptoms.

Importance of Early Identification and Obstacles to Early Diagnosis

The ability accurately to identify and differentiate among presenting problems (including identification of the presence of comorbid conditions) has important implications for later outcomes. 'Pure' diagnoses differ from comorbid disorders with regard to symptomology, outcomes, and treatment responsiveness (Kadesjo et al., 2003; Kaplan et al., 2001). Research on the developmental trajectory of ADHD, ODD, and comorbid ADHD+ODD suggests differing developmental pathways (Waschbusch, 2002), different types of associated problems, and varying outcomes (Cunningham & Boyle, 2002; Keenan & Wakschlag, 2002).

Differential diagnosis and the identification of comorbid disorders are particularly important for the development of appropriate and effective interventions (Connor, 2002). In a study of clinic-referred children (ages 7 to 10) who met diagnostic criteria for ADHD, Jensen and

colleagues (2001) found that treatment response differed across children with various diagnostic profiles. After being divided into four diagnostic groups (ADHD-only; ADHD+Anxiety; ADHD+ODD/CD; ADHD+Anxiety+ODD/CD), participants were randomly assigned to one of four treatment conditions: (a) medication treatment, (b) behavioral intervention, (c) medication treatment and behavioral intervention, (d) standard community treatment. Participants with ADHD and those with ADHD+ODD/CD appeared to respond best to medication treatment. In contrast, children with ADHD and comorbid anxiety responded better to behavioral intervention or the combination treatment, depending on the outcome measure used to evaluate treatment response. Although responsiveness to treatment varied across different outcome measures, the findings suggest that diagnostic profile may have important implications for treatment selection.

The large-scale MTA study produced similar results. Biederman and colleagues (2006) reported that the presence of comorbid ODD did not impact the degree to which ADHD symptoms were improved by treatment. However, ODD symptoms demonstrated the greatest improvement in response to medication treatment, rather than behavioral intervention. Thus, it appears that early identification of individuals at risk for psychiatric disorders and differentiation between their diagnostic profiles could contribute significantly to the selection of appropriate interventions.

Unfortunately, there are significant obstacles to effective early diagnosis. Efforts toward early identification and intervention have revealed that although there is substantial evidence that NDs and DBDs emerge during the preschool years (Lavigne et al., 1996; Egger & Angold, 2006; Keenan & Wakschlag, 2000; Wilens et al., 2002) and are relatively stable over time (Campbell et al., 1984; Campbell, Breaux, Ewing & Szumowski, 1986a, Lahey et al., 2004; Lee et al., 2008; Speltz et al., 1999), there are limitations to the relevance of DSM criteria for preschool-aged

children. Many of the symptoms of neurodevelopmental and disruptive behavior disorders are common among typically-developing preschoolers (Connor, 2002; Mahone & Schneider, 2012), making it difficult to distinguish behavior that is developmentally appropriate from behavior that warrants further evaluation/intervention. Keenan and Wakschlag (2000) cited the failure of the DSM to include "developmentally specific anchors" as a significant obstacle to diagnosis at the preschool level (p. 34). Other researchers have noted that the diagnostic criteria for DBDs were developed for older children (Arons et al., 2002; Thorell & Wählstedt, 2006). Several factors make it difficult to apply the existing diagnostic criteria effectively to preschool-aged children, including rapid changes in developmental abilities, variations in environmental demands, and the difficulty of distinguishing transient behavior problems from those that persist over time and may require intervention (Banaschewski, 2010).

Dimensional Approach to Diagnosis

Criticisms of a categorical approach to diagnosis have led many researchers to consider a dimensional approach, in which behaviors are conceptualized in terms of their placement on a continuum of severity, rather than as distinct symptoms (Lavigne et al., 1996; Sonuga-Barke et al., 2005). In a dimensional approach, information is gathered not only about the presence and number of problem behaviors, but also about the nature, frequency, and severity of the concerns. Within a dimensional framework, assessment typically involves administration of behavior rating scales and the use of cut-off scores to identify individuals whose behavior is significantly different from the norm (Egger & Angold, 2006; Jensen et al., 1996). Behaviors falling on the severe end of the continuum are considered risk factors rather than conclusive evidence of a disorder (Egger & Angold, 2006).

There are numerous advantages to a dimensional approach to diagnosis. Because it involves

consideration of more information about behavior (e.g., frequency, intensity, normative comparison) than a categorical method of identification, the dimensional approach may be more accurate in identifying individuals at risk for adverse outcomes. This notion is supported by findings of substantial impairment among individuals who do not meet full diagnostic criteria (Egger & Angold, 2006; Fergusson and Horwood, 1995). Diagnostic decisions based on a dimensional approach to assessment may also be more reliable (Achenbach et al., 2008), given their utilization of a more structured method of evaluation, and reliance on empirical data regarding the distinction between normative and deviant patterns of behavior.

The validity of the dimensional approach to assessment is supported by its alignment with the apparent underlying structure of various symptoms and disorders, which have been described as more dimensional than categorical in nature (Marcus & Barry, 2011). In particular, several studies of the manifestation of symptoms of ADHD over time have demonstrated instability in subtype classifications (Lahey et al., 2005; Lahey & Willcutt, 2010), while indicating that the majority of children who meet diagnostic criteria early on will continue to do so, albeit under a different subtype. This has led researchers to suggest the use of a dimensional approach to the diagnosis of ADHD, based on the number of symptoms within a given domain (inattentive versus hyperactive/impulsive) that are present at the time of assessment, rather than persisting in the use of nominal classification (Willcutt et al., 2012).

In addition to the general advantages associated with dimensional diagnosis, this approach may be particularly well-suited for use with young children. Given that many of the behaviors symptomatic of NDs and DBDs are prevalent among preschoolers, the presence of a psychiatric disorder may be distinguished from normative behavior only by such features as frequency, chronicity, and severity (Kadesjo et al., 2003). Placing assessment within a normative context is

also important, given the rapid, qualitative changes that occur in young children's cognitive, social, and behavioral functioning (Arons et al., 2002).

Although use of a dimensional, more empirically-based approach to diagnosis could be beneficial in addressing various disadvantages of the more prominent categorical approach, it is important to note several limitations to this approach. Selection of appropriate assessment instruments is one obstacle to a dimensional approach to assessment and diagnosis. Interviews and rating scales are common methods of assessment (Ryan-Krause, 2010), but there is a lack of consensus regarding the specific diagnostic tool(s) or approach that should be used (Bussing et al., 2006; Keenan et al., 1997; Smith & Corkum, 2007). Lack of consensus regarding selection of a cut-off score to differentiate typical from problematic behavior levels also complicates the effective use of behavioral assessments (Serna et al., 2002). The establishment of a cut-off score requires the determination of "acceptable" levels of false positive and false negative decisions, which likely differ across clinicians. Finally, although there is general agreement that assessment should involve input from multiple respondents across various settings (Mahone & Schneider, 2012; Ryan-Krause, 2010; Skovgaard, Houmann, Landorph & Christiansen, 2004), low correspondence in ratings across informants (Drabick et al., 2004) leads to questions regarding how to interpret inconsistencies when ratings differ.

An important step in advancing the knowledge base would involve comparing the results of dimensional and categorical approaches. Such analysis could provide important information regarding the relative strengths and weaknesses of each approach. It is also important to determine whether the two approaches produce similar outcomes in terms of the diagnostic status of various individuals. The most comprehensive and meaningful approach to diagnosis could involve the combined use of both dimensional and categorical methods.

Comparisons between Diagnostic Approaches

Several studies have examined the relationship between categorical DSM diagnoses and dimensional approaches to assessment/diagnosis among children and adolescents who met criteria for various psychiatric disorders. Jensen and colleagues (1993) examined the relationships between parent- and child- diagnostic interviews (using the DISC, a categorical approach) and parent-completed rating scales (a dimensional measure) for children ages 5 to 17. The rating scales used in analysis included the CBCL, the Children's Depression Inventory (CDI), and the Revised Children's Manifest Anxiety Scale (RCMAS). Because it is more relevant to the current topic, only the results pertaining to the CBCL are reviewed here.

The CBCL (as completed by parents) was found to have some utility in differentiating among individuals with various externalizing diagnoses, but was less effective at identifying internalizing disorders (Jensen et al., 1993). The Anxiety and Depression subscales of the CBCL were most elevated among participants with a diagnosis of Conduct Disorder (CD), rather than diagnoses of anxiety or depression, as would be anticipated. In addition, pairwise comparisons of the various DISC-based diagnoses and CBCL subscales did not reflect a significant relationship between DSM diagnosis and the anxiety or the depression subscale. Participants with ODD and CD diagnoses received higher ratings (reflecting increased concern) on the Aggression and Delinquency subscales of the CBCL (respectively); however, scores on these subscales were also elevated among participants with other diagnoses (e.g., ADHD, depression). Similar results occurred for the Hyperactivity subscale of the CBCL, with individuals with CD and ODD receiving the highest scores, but individuals with other diagnoses receiving only "somewhat lower" scores (Jensen et al., 1993, p. 405). Although the analyses produced several unexpected results, the findings suggest that the CBCL may be of utility in identifying individuals with

externalizing disorders. The degree of elevation in scores appeared to be a key factor differentiating among individuals with various diagnoses. It is also important to note that the results may have been affected by the authors' inclusion of a limited number of CBCL scales. In addition, many of the participants had both internalizing and externalizing disorders and/or symptoms that fell below traditional diagnostic thresholds, complicating the task of differentiating among symptom patterns. The presence of comorbid disorders adds to the complexity of determining the presence and nature of the relationship(s) between dimensional measures of behavior and a categorical diagnosis.

Research examining singular (non-comorbid) DSM diagnoses has suggested a correspondence between categorical and dimensional approaches to identification. In a large-scale investigation of the accuracy of the CBCL in predicting ADHD diagnosis, Chen and colleagues (1994) compared rating scale scores of 140 children with ADHD and 120 comparison children from a pediatric setting. The participants were between the ages of 6 and 18. Diagnosis was based on parent and child diagnostic interviews using the Schedule for Affective Disorders and Schizophrenia (SADS). Results of the analyses (focused only on identification of an ADHD diagnosis) revealed that the Attention Problems subscale of the CBCL significantly discriminated between children with and without ADHD. Addition of other subscales did not significantly contribute to diagnostic prediction. These results were replicated in a sample of siblings of the original participants (174 of siblings of the children with ADHD and 129 siblings of the pediatric sample).

The results of these studies suggest that various dimensional measures may have utility in the identification of neurodevelopmental and disruptive behavior disorders among children and adolescents across a relatively broad age range; research conducted exclusively with preschool-

aged children has yielded similar results. Keenan and Wakschlag (2000) examined the relationship between DSM-IV diagnosis of ODD or CD and behavior ratings and observations. Diagnoses were identified on the basis of parent interview, using the fifth version of the Schedule for Affective Disorders for School-Age Children – Epidemiological (K-SADS). The participants were 79 children (between the ages of two and five) consecutively referred to a clinic for preschoolers with behavior problems.

The majority of children with an Externalizing T-score over 70 met criteria for CD (59%) or ODD (81%; Keenan & Wakschlag, 2000). The converse was also true: of children who met criteria for a diagnosis of ODD or CD, the majority had an elevated T-score (above 70) on the Externalizing scale of the CBCL (66% of those with CD, 62% of those with ODD). Rating scale scores did not distinguish between children with ODD and those with CD.

Direct observation revealed that noncompliance was relatively common, exhibited by 20% of the participants during a puzzle completion task, and by 29% of the participants during a clean-up task (Keenan & Wakschlag, 2000). Aggression was less common, exhibited by 2.5% and 3.8% of participants during the clean-up and puzzle completion tasks, respectively. Observed levels of noncompliance and aggression during the puzzle completion task were significantly associated with a CD diagnosis, but not with an ODD diagnosis.

These results suggest that among clinic-referred preschoolers, there was a correspondence between dimensional measures and ODD and CD diagnoses as defined through the traditional, categorical approach (Keenan & Wakschlag, 2000). However, interpretation of these results should be placed within the context of the study's limitations, including a small sample size (N = 79) and the relatively homogeneous nature of the sample: clinic-referred children who were predominantly African American (over 80%) males (77%). Interestingly, the authors noted that

several modifications were made to the DSM diagnostic criteria to render them more appropriate for a preschool population, such as elimination of the requirement that the symptom of stealing involve an item of some value, and addition of the requirement that temper tantrums be of heightened frequency and duration in order to qualify as a symptom.

Overall, the few studies that have analyzed the relationship between behavior ratings and categorically defined psychiatric diagnoses provide support for a relationship between various CBCL subscales (e.g., externalizing, attention problems, and hyperactivity) and neurodevelopmental and disruptive behavior disorders as defined in the DSM. However, the current knowledge base remains limited in numerous ways. Most of the extant studies have been conducted with school-aged children and/or adolescents, the majority relying solely on a single informant (parent ratings) and/or clinic-based measures to evaluate child characteristics, despite evidence that the use of multiple sources of information enhances both the sensitivity and specificity of diagnostic decisions (Sprafkin, Gadow & Nolan, 2001). In addition, the research has focused primarily on examining the correlation between rating scale scores (or other behavioral measures) and diagnosis, without evaluating the combined utility of categorical and dimensional methods for distinguishing among diagnoses and/or for identifying the presence of comorbid disorders. The current study will expand the current knowledge base by focusing exclusively on a preschool population and by examining the relationships among multiple measures of behavior and skills and their degree of correspondence with categorically defined disorders.

Chapter III: Methods

Participants

This study included 70 preschoolers (Table 1) between the ages of 37 and 71 months (M = 52.23; SD = 8.40). The participants were predominantly male (74.3%) and Caucasian (72.9%), although other ethnic groups were represented: Hispanic (11.4%), African-American (1.4%). Ten of the participants (14.3%) were of an ethnicity not listed among the options provided ("other"). As part of the screening measures (further described below), the DBD portion of the Diagnostic Interview Schedule for Children, Version IV was administered to parents via telephone interview (DISC-IV; Shaffer, Fisher, Lucas, Dulcan & Schwab-Stone, 2000). The strong majority of participants met criteria for comorbid ADHD and ODD (N = 55, 78.6%). Two of the participants (2.9%) met criteria for the Inattentive Type of ADHD only; seven (10.0%) met criteria for ADHD, Hyperactive-Impulsive Type only; and six (8.6%) met criteria for ADHD, Combined Type only. Because inclusion criteria required that the participants exhibit characteristics of ADHD, none of the participants met criteria for a diagnosis of ODD alone (0.0%).

Recruitment

The current sample was derived from a larger sample of preschoolers (N = 135) who participated in a five-year longitudinal study examining the comparative effectiveness of two psychosocial interventions for ADHD (Kern et al., 2007). Recruitment for the longitudinal study occurred through pediatricians' offices, preschools, and day care centers. Information about the project was periodically disseminated to these facilities via brochures and follow-up phone calls. Parents who expressed an interest in participation were screened through telephone interview to determine the presence of ADHD symptoms and to ascertain attendance at a preschool or daycare at least two days per week. Preschoolers who met these criteria were included in a three-

stage screening process.

Screening Measures. During the first stage of screening, the Conners' Parent and Teacher Rating Scales – Revised (Conners, 1997) were administered to ascertain the presence of ADHD symptoms. The rating forms were hand-delivered or mailed to parents and teachers for completion. Inclusion criteria required a T-score of 65 or above (1.5 standard deviations above the mean; 93rd percentile) on both the parent- and teacher-report versions for at least one ADHD-related subscale. Preschoolers who met these criteria advanced to the second stage of screening. Although the Conners rating forms were used only as a screening measure for inclusion in the larger, longitudinal study, they were among the key variables included in analysis in the current study.

Subsequent stages of screening involved measures designed to exclude children with Autism and/or Conduct Disorder, those with an overall cognitive ability below a Standard Score of 80 (due to the possible presence of a developmental disability), as well as children whose reported symptoms did not produce functional impairment (as measured by the Children's Global Assessment Scale (CGAS; Shaffer et. al., 1983). Further assessment of ADHD and ODD symptoms was conducted via a structured telephone interview with parents, which included administration of the DBD portion of the DISC-IV. Inclusion criteria required that children meet diagnostic criteria for at least one of the three subtypes of ADHD (as assessed by the DISC-IV).

Of the 536 preschoolers referred for participation in the longitudinal study, inclusion criteria (including parental consent to participate) were met for 152. Data from 17 children were subsequently excluded, due to scoring errors or psychiatric conditions that were subsequently discovered. Data were reported for a total of 135 participants (Kern et al., 2007).

All participants with a complete baseline dataset were included in the current study, resulting

in a sample of 70 participants. Chi square analyses were conducted to determine whether there were statistically significant differences between participants included in the current study and the full sample included in the longitudinal study, with regard to relevant demographic characteristics (Table 2). It is important to note that the assumption regarding the minimum expected frequency in each cell was violated for all of the analyses, with the exception of comparisons of participant age. For this reason, the results should be interpreted with caution. The results did not indicate any statistically significant differences between the longitudinal sample and the subsample included in the current study with regard to gender, ethnicity, parents' level of education, medication status (whether the child was taking a psychotropic medication), or the number of days per week the child attended an educational program/activity. A statistically significant difference was evident between the two samples with regard to participant age ($p < 0.05$). The participants in the current sample were somewhat younger than those in the longitudinal study, including more three-year-olds and fewer four- and five-year-olds.

Measures

The measures included in the current study were administered to all participants at baseline. Various assessments of attention and behavioral and social functioning were selected for analysis, including parent- and teacher- completed rating scales as well as direct observation measures. All rating forms were either hand-delivered or mailed to parents and teachers for completion. Due to the longitudinal nature of the study for which the participants were recruited, several of the measures used in assessment were developed for, and normed on, older (school-aged) children to ensure continuity in assessment methods over time. Thus, the measurement tools used in the current study may not reflect best practice for assessment of preschool children in a clinical setting, for diagnostic purposes. Because standard scores were not available for

measures normed on an older population, raw scores were used for all variables included in analysis to maintain consistency in scoring across measures.

Child Behavior Checklist and Teacher Report Form (Achenbach, 1991a; Achenbach, 1991b). The Child Behavior Checklist (CBCL) and Teacher Report Form (TRF) are broad-band measures of behavioral functioning. The CBCL, used for individuals between the ages of 4 and 18 years, is a parent rating form comprised of 118 items. The TRF (comprised of 113 items) is completed by teachers for individuals between the ages of 5 and 18 years. Both measures present various statements about the child's behaviors and/or emotions which are rated on a three-point Likert scale: 0 = "not true," 1 = "somewhat or sometimes true," 2 = "very true or often true."

Scoring of the items yields three global scales: Total Problems, Internalizing, and Externalizing. Scores are also produced for eight syndrome subscales, derived via principal components analysis. The Internalizing Scale includes three subscales: Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints. The two subscales comprising the Externalizing Scale are: Delinquent Behavior and Aggressive Behavior. The remaining three subscales did not consistently correspond with the Internalizing or Externalizing scales and are considered independently: Social Problems, Thought Problems, and Attention Problems (Achenbach, 1991b). For the current study, global scales were not included in analysis. Raw scores for all of the CBCL and TRF subscales were initially included in analysis.

The CBCL and TRF are among the most widely used measures of child behavior problems and competencies (Biederman et al., 1993). When administered to children ages 6 – 18 years old, the test-retest reliability of the CBCL (based on a 1-week interval between assessments) was found to be strong, ranging from 0.60 (Activities subscale for girls) to 0.97 (Somatic Complaints subscale for boys). The majority of test-retest *r* values were at or above 0.80 (32 of 36

correlation values; Achenbach, 1991b).

Assessment of the inter-rater reliability of the CBCL was based on evaluation of the correspondence between the ratings provided by children's mothers and fathers (to prevent confounding test-retest reliability with inter-rater reliability). Inter-rater reliability ranged from 0.26 (Thought Problems subscale for girls aged 12-18) to 0.91 (School subscale for boys aged 12-18). When collapsed across age range and gender, the average *r* values for inter-rater reliability ranged from 0.48 to 0.87, most falling above 0.70 (Achenbach, 1991b).

The internal consistency of the CBCL was evaluated via Cronbach's *alpha*. The resulting values were reported in the user's manual (Achenbach, 1991b). Cronbach's *alpha* values for the various subscales ranged from 0.42 to 0.96, with most falling within acceptable range.

Criterion validity of the CBCL and TRF was evaluated with regard to correspondence between rating scale scores and referral status (Achenbach, 1991b). The rating scale scores of demographically matched participants were analyzed for significant differences between the scores of non-referred children and those of children referred for mental health services (for the CBCL; Achenbach, 1991b) or for special education services due to emotional/behavioral problems (for the TRF; Achenbach, 1991a). Regression analysis was used to determine the percentage of variance in TRF or CBCL score accounted for by referral status. For both rating scales, the results revealed a consistent pattern of lower scores on the adaptive scales and higher scores on the problem scales among participants referred for services, supporting the validity of the CBCL and the TRF in identifying individuals with problematic behavior patterns.

Conners' Rating Scales – Revised: Long Form (Conners, 1997). The Conners' Rating Scales – Revised are measures designed to gather information about attentional, emotional, and behavioral problems. The parent-report and teacher-report forms are comprised of 80 and 59

items (respectively), with responses presented on a four-point Likert scale ranging from 0 ("not at all true") to 3 ("very much true"). The Conners' Teacher Rating Scale – Revised: Long Form (CTRS-R:L) yields six primary subscales: Oppositional, Hyperactivity-Impulsivity, Cognitive Problems/Inattention, Social Problems, Anxious/Shy, and Perfectionism. The parent rating form (CPRS-R:L) yields the same primary subscales as well as a Psychosomatic subscale. In addition to these core subscales, the Conners Rating Scales also yield five broader measures, derived from these subscales: the Conners' Global Index, three DSM-IV Symptoms Indices (Inattention, Hyperactivity, and Impulsivity), and the ADHD Index. For the current study, raw scores for all of the CTRS-R:L and CPRS-R:L subscales were initially included in analysis.

Examination of the CPRS-R:L was conducted on a sample of 2,200 children between the ages of 3 and 17 years (Conners, 1997). Among individuals aged three to seven years, the internal reliability of the CPRS-R:L was found to be strong, with coefficient alphas ranging from 0.77 to 0.92. From the larger sample, a subset of 49 children was rated twice by a parent, approximately six weeks apart to evaluate the test-retest reliability of the scale. The results revealed moderate to strong test-retest reliability, with nearly all test-retest correlations ranging from 0.42 to 0.78 (results were not reported by age range). The only exception occurred on the Social Problems subscale, for which the test-retest correlation was small (0.14) and non-significant.

Criterion validity of the CPRS-R:L was evaluated through comparison of parent ratings of clinic-referred children with an independent diagnosis of ADHD with ratings of children without ADHD, randomly selected and matched for age, sex, and ethnicity (Conners, Sitarenios, Parker & Epstein, 1998a). Ratings of the children with a diagnosis of ADHD were significantly more elevated than the ratings of the control comparisons on six of the seven subscales: Oppositional,

Hyperactivity-Impulsivity, Cognitive Problems, Social Problems, Psychosomatic, and Anxious-Shy. The only exception was the Perfectionism scale, on which parent ratings did not reflect a significant difference between children with ADHD diagnosis and control comparisons. When discriminant function analysis was used to predict children's membership in one of two groups (with ADHD or without ADHD), the CPRS-R:L was found to have 92% sensitivity and 94% specificity. These results provide support for the criterion validity of the CPRS-R:L.

Examination of the CTRS-R:L was conducted on a sample of 1,702 children between the ages of 3 and 17 years (Conners, Sitarenios, Parker & Epstein, 1998b). The internal reliability of the CTRS-R:L for children aged three to seven was found to be strong, with coefficient alphas ranging from 0.88 to 0.95. From the larger sample, 50 children were randomly selected for repeated ratings by teachers to evaluate the test-retest reliability of the scale. The results revealed moderate to strong test-retest reliability, with test-retest correlations ranging from 0.47 to 0.88 (aside from the Hyperactivity-Impulsivity subscale, the test-retest correlations for all other subscales were above 0.60; results were not reported by age range).

Criterion validity of the CTRS-R:L was evaluated through comparison of teacher ratings of clinic-referred children with an independent diagnosis of ADHD with ratings of children without ADHD, randomly selected and matched for age, sex, and ethnicity (Conners et al., 1998b). Ratings of the clinic-referred children were significantly more elevated than the ratings of the control comparisons on the following scales: Oppositional, Hyperactivity-Impulsivity, Inattention/Cognitive Problems, Social Problems, and Anxious-Shy. In contrast, the children without ADHD had higher ratings on the Perfectionism scale. When discriminant function analysis was used to predict children's membership in one of two groups (with ADHD or without ADHD), the CTRS-R:L was found to have 78% sensitivity and 92% specificity. These

results provide support for the criterion validity of the CTRS-R:L.

Social Skills Rating System, Elementary Version: Parent and Teacher Forms (Gresham & Elliott, 1990). The Social Skills Rating System (SSRS) rating scales provide a standardized measurement of social behaviors, as rated by various informants (Gresham & Elliott, 1990). Items are rated on a 3-point Likert scale, based on the frequency of various behaviors (0 = "Never," 1 = "Sometimes," 2 ="Very Often"). The current study employed the Elementary versions the 55-item parent rating form (SSRS-P) and the 57-item teacher rating form (SSRS-T). The SSRS-P and SSRS-T were designed for assessment of children in Kindergarten through sixth grade. Both measures yield standard scores in two key domains of functioning: Social Skills and Problem Behaviors. Although the SSRS-T also includes an Academic Competence scale, due to the current focus on behavioral functioning, only (raw) scores on the SSRS-P and SSRS-T Social Skills and Problem Behaviors scales were included in the current analysis.

The psychometric properties of the teacher-rated and parent-rated forms of the SSRS were examined in large, nationally representative samples (Gresham & Elliott, 1990). A group of 259 teachers provided ratings for 1,335 children. Ratings were also provided by 1,023 parents or guardians (for 1,023 children). For children at the preschool level, internal consistency values ranged from 0.57 (Internalizing subscale, Parent Form) to 0.94 (Total Scale, Teacher Form). At the elementary level, internal consistency values ranged from 0.65 (Responsibility, Parent Form) to 0.94 (Total Scale, Teacher Form).

Test-retest reliability of the SSRS was conducted over a 4-week period (Gresham & Elliott, 1990). Reliability coefficient values were reported only for children at the elementary level. Teacher ratings yielded strong test-retest reliability coefficients, ranging from 0.75 (Assertion) to 0.93 (Academic Competence). The results of parent ratings yielded test-retest reliability

coefficients ranging from 0.48 (Internalizing) to 0.87 (Total Scale).

Content validity of the SSRS is supported by the reliance on empirical literature in the development of the items and by the use of Importance ratings for each item (Gresham & Elliott, 1990). Evidence of criterion validity is reported in terms of the correspondence between teacher ratings on the SSRS-T and the Social Behavior Assessment (SBA) as well as the CBCL. Correlations between subscales of the SSRS-T and the SBA, reported for children at the elementary level, ranged from -0.01 (SSRS-T Internalizing and SBA Environmental subscales) to 0.73 (SSRS-T Cooperation and SBA Task-Related subscales). Correlations between subscales of the SSRS-T and the CBCL ranged from 0.11 (SSRS-T Externalizing and CBCL Internalizing) to 0.81 (SSRS-T Total Scale and CBCL Externalizing scale).

Evidence of the criterion validity of the SSRS was also reported in terms of the correspondence between parent ratings on the SSRS-P and on the CBCL (Gresham & Elliott, 1990). Correlation values ranged from 0.00 (SSRS-P Cooperation and CBCL Activities subtests) to 0.74 (SSRS-P Hyperactivity and CBCL Externalizing subscales). It is important to note that the measures used in evaluation of the criterion validity of the SSRS assess a broad range of skills and problem behaviors. A strong correlation between divergent features of behavior would not be expected.

Direct Observation. In addition to the rating scales described above, data regarding participants' behavior were also gathered via direct observation in the preschool/daycare setting. The Classroom Observation Code (COC; Abikoff & Gittelman, 1985) was used to record behavior during structured activities (such as circle time discussions and academic instruction). This structured observation system was designed to provide information relevant for diagnostic decisions with regard to ADHD (Volpe, DiPerna, Hintze & Shapiro, 2005) and involves

recording of the occurrence of 12 mutually exclusive behaviors, gathered in 15-second intervals. Most of the behaviors are recorded on a partial interval schedule: (a) interference (e.g., talking or playing around during work-time); (b) interference to the teacher (e.g., interrupting lecture); (c) minor motor movement (e.g., fidgeting while seated); (d) gross motor standing (e.g., out of seat and standing); (e) gross motor vigorous (e.g., running); (f) physical aggression; (g) verbal aggression toward the teacher; (h) verbal aggression toward a peer; and (i) solicitation of the teacher (e.g., raising hand, seeking attention). The remaining three behaviors are recorded via whole interval method: (a) off-task; (b) non-compliance; (c) out of chair. If none of these behaviors occurs during an observed interval, "absence of behavior" is recorded. Scoring of the COC yields a rate of each behavior (Abikoff et al., 2002).

The reliability and validity of the COC were evaluated by Abikoff and colleagues (1977) based on classroom observations of 120 children aged 6 to 12 years. The participants included 60 children referred by schools, based on parent request for treatment due to concerns about hyperactivity and significant behavior problems. These participants were paired with 60 non-referred children of a similar age, matched for gender. Each observation was 32 minutes long; every four minutes, the observer rotated between recording the behavior of one of the target (referred) children and one of the control comparisons.

Inter-observer reliability was calculated via phi coefficients (Abikoff, Gittelman-Klein & Klein, 1977). The average phi coefficient was 0.76 across all categories of observed behavior (range: 0.24 to 1.00). The overall reliability of each observation session was calculated using the product-moment correlation between the total occurrences of behavior as recorded by the "standard observer" and by the secondary observer, across the entire observation period (Abikoff et al., 1977, p. 777). The majority of the correlations were above 0.90 (range: 0.14 to 1.00). The

behaviors that occurred most frequently and which were observed with the greatest inter-observer reliability were: interference, off-task, minor motor movement, gross motor movement, and solicitation of teacher attention.

The validity of the COC observation method was evaluated by comparing the behavior of target children (referred due to concerns regarding hyperactivity and behavior problems) with the behavior of the comparison controls. Correlated *t* tests revealed that the mean behavior scores of the target children were significantly more elevated than the mean scores of the comparison children, indicating that the data gathered via the COC accurately distinguished children identified as having behavior problems from typical peers. Further analysis revealed that the behavior of the hyperactive children was more variable, while the control comparisons exhibited behavior that was unimodal and positively skewed, with low levels of most of the target behaviors.

In a subsequent study, Abikoff and colleagues (1980) cross-validated the previous findings in another sample of children aged 6 to 12 years. Most of the participants in the second study were from different schools; 61 referred children met inclusion criteria and were paired with 61 control comparisons matched for gender and approximate age. Inter-observer reliability was again calculated via phi coefficients. The mean phi coefficient value across all observers and all behaviors recorded was 0.82 (range: 0.01 to 1.00).

Overall, the results of the replication study were highly consistent with those of the original study. Children referred for hyperactivity and behavior problems exhibited significantly more instances of the behaviors targeted for observation than comparison controls (Abikoff, Gittelman & Klein, 1980). In both studies, the behaviors that best discriminated the hyperactive children from their matched comparisons were off-task behavior and interference. The results indicated

that hyperactive children were most accurately identified through examination of these two target behaviors.

The current study included COC observation data gathered in the preschool or daycare setting during large-group, structured activities. Each participant was observed on one occasion (at baseline), for approximately 20 minutes. The number of intervals each participant was observed ranged from 41 to 60. In order to control for this variation, rates were calculated for each behavior by dividing the number of behaviors recorded as occurring by the number of intervals observed. Data regarding the following behaviors were initially included in analysis: interference, off-task, noncompliance, minor motor movement, gross motor standing, out of chair, physical aggression, verbal aggression/threat toward a child, verbal aggression/threat toward a teacher, gross motor vigorous, solicitation of teacher, absence of behavior, teacher approval, and teacher disapproval.

The Early Screening Project (Feil, Severson & Walker, 1998) observation system was used to assess behavior during unstructured activities (e.g., free play). The Early Screening Project (ESP) was designed as a three-stage screening process for identifying behavior disorders among preschool-aged children. At each stage, the criteria for proceeding to the next stage are based on "at risk" status, as defined by deviating from the mean by at least one standard deviation. The first stage involves teacher rankings of children with regard to internalizing and externalizing behavior. Children ranked as among the top three with regard to internalizing behaviors and those ranked among the top three with regard to externalizing behaviors progress to the second stage. In Stage 2, teachers complete ratings of children on four behavioral measures assessing: critical events, aggressive behavior, adaptive behavior, and maladaptive behavior. Children whose scores in each behavioral domain meet or exceed established cut-offs progress to the third

stage. Due to behavioral differences between boys and girls, the cut-off scores for two of the four measures (aggressive behavior and adaptive behavior) differ by gender. Stage 3 consists of direct behavioral observation of children's social behavior during free play situations in the classroom or on the playground. Children's behavior is coded as being prosocial or antisocial/nonsocial. Behaviors recorded as antisocial/nonsocial include: negative verbal or physical interactions, disobeying classroom rules, tantruming, or engaging in solitary play (nonsocial behavior). Stopwatches are used to tally the total proportion of time children spent in positive or negative/nonsocial behavior. The cut-off for identifying children as "at risk" is based on the percentage of time they spent engaged in aggressive or solitary play. Again, the specific cut-off criteria differ for boys (≥ 40% of the observation time) versus girls (≥ 37% of the observation time).

The reliability and validity of an early form of the Preschool Screening for Behavior Problems (PSBP) were evaluated by Feil and Becker (1993). The screening process was applied to 105 preschoolers between the ages of three and six. Observations were conducted to gather data about the children's behavior in two categories: engaged time during structured activities and peer/social behavior in unstructured settings. Two 10-minute observations were completed during each activity type, for a total of 40 minutes of observed time. Target behaviors during observation of engaged time included attending to the teacher, following directions, and requesting help appropriately. Children's behavior during the peer social behavior observations was coded according to several categories. Social engagement and participation were coded as being positive or negative in nature. Additional behavioral categories were coded for the presence or absence of various target behaviors, including: "parallel play, solitary play, alone, adult, and no codeable response" (Feil & Becker, 1993, p. 46). For the purposes of further data

analysis, observational data were aggregated across sessions by behavior category. The behavioral categories with the highest correlations with other measures (Stage 2 screening ratings, Behar rating scale scores (Behar, 1977), and Conners' Hyperactivity and Inattention scales) were combined to produce two general behavior categories: positive peer social behavior and negative peer social behavior.

Observers received training until they reached 90% accuracy with behavioral coding of videotaped scenes. Subsequently, inter-rater reliability was calculated for 20% of randomly selected observational sessions. Inter-rater reliability for direct observation of engaged time during structured activities was calculated by dividing the sum of the smaller scores across children by the larger scores, producing a "percentage indicator of rater differences weighted for length of observation" (Feil & Becker, 1993, p. 49). This calculation resulted in a reliability coefficient of 0.97. Inter-rater reliability for observations of peer social behavior was calculated according to interval-by-interval agreement between observers. This calculation yielded an average agreement rate of 0.87. Overall, inter-rater reliability of the observation system was supported.

Test-retest reliability of the PSBP was assessed through comparison of ratings and observations completed in November and December with those completed in March and May (Feil & Becker, 1993). Although there were statistically significant correlations between children's observed behavior across observations, the magnitude of the relationships was small (Pearson *r* values ranged from 0.20 for engagement during structured activities to 0.22 for positive and negative peer social behavior). However, the authors noted that due to the situational dependency of preschoolers' behavior, low test-retest correlations were expected.

Discriminant validity of the PSPB was evaluated via comparison across three groups of

children: (a) those ranked highest by teachers with regard to externalizing behavior, (b) those ranked highest by teachers with regard to internalizing behavior and (c) a comparison group of children consistently rated by teachers as average across PSPB measures as well as the Behar and Conners' Rating Scale across both the fall and spring assessments (Feil & Becker, 1993). The mean rates of behavior exhibited by each group remained generally consistent from the fall to the spring assessment, reflecting some consistency over time when analyzed at the group level. Although the observation data alone did not clearly distinguish among the three groups, certain behavior patterns emerged. Children in the externalizing and internalizing groups both had similarly low levels of prosocial behavior, suggesting social skill deficits. Children in the internalizing group had higher levels of engaged time during structured activities than children in the externalizing group or control comparisons.

The validity of the ESP as a screening method was examined by Feil and colleagues (1998) through comparison of referred and non-referred children on several behavioral measures. The results of *t*-tests revealed significant differences between the observed behavior of children who met criteria for "at risk" status in Stage 3 (the observation portion) of the ESP process and those who did not meet Stage 3 criteria and therefore would not have been referred. Scores on the Conners' Hyperactive and Inattentive subscales and on the CBCL Aggression and Withdrawal subscales were higher (reflecting increased concern) among children who met ESP criteria for referral status compared with those who fell below referral criteria. These results provided support for the utility of the ESP with regard to identifying children with behavioral problems.

The current study included ESP observation data gathered in the preschool/daycare setting during unstructured activities. Each participant was observed on one occasion (at baseline), for approximately 20 minutes. The number of intervals each participant was observed ranged from

36 to 80. In order to control for this variation, rates were calculated for each behavior by dividing the number of behaviors recorded as occurring by the number of intervals observed. Data regarding the following behaviors were initially included in analysis: negative verbal interactions, negative physical interactions, disruptive behavior, off-task behavior, activity change, solitary play, parallel play, and positive social engagement.

Data Collection Procedures

The measures included in the current study were gathered at baseline, upon entry into the longitudinal study of which these data are a subset. As noted previously, the rating scales were either mailed or hand-delivered to parents and teachers (accompanied by a self-addressed, stamped envelope). Direct observations were conducted by graduate students in a school psychology or special education program. Observers were blind to the purposes of the study and to participants' group membership. Rating scales were scored by data collectors; accuracy was assessed via independent hand scoring of 30% of assessments. Accuracy of data entry within the master database was also evaluated for 30% of assessments via direct comparison of raw data to each entered item. Very few errors were evident in scoring or data entry.

For direct observation data, interobserver agreement was calculated for 40% of randomly selected observations across participants. Interobserver agreement (IOA) for the occurrence and non-occurrence of each behavior was calculated on an interval-by-interval basis. For observations using the Abikoff COC system (Table 3), the average IOA value regarding the occurrence of the target behaviors (across all behaviors included in the current study) was 83.6% (range: 70% - 100%). The average IOA value regarding the non-occurrence of the target behaviors (across all behaviors included in the current study) was 98.4% (range: 91% - 100%). The mean kappa was 0.86 (range: 0.75 – 1.0).

For observations using the ESP method (Table 4), the average IOA value regarding the occurrence of the target behaviors (across all behaviors included in the current study) was 81.7% (range: 71% - 90%). The average IOA value regarding the non-occurrence of the target behaviors (across all behaviors included in the current study) was 94.3% (range 85% - 100%). The mean kappa was 0.83 (range: 0.76 – 0.91).

Data Analysis

Preliminary Data Analysis: Multicollinearity. In order to prevent undue weighting of constructs due to strong correlations between variables (Sambandam, 2003), the data were evaluated for multicollinearity. Statistically significant correlations that fell within the moderate to strong range ($r \geq \left| 0.60 \right|$) were examined for potential redundancy among variables. Variables determined to be redundant were removed from analysis.

Preliminary Data Analysis: Variable Scaling. The inclusion of variables with different scaling can present a problem in cluster analysis, as variables with larger standard deviations are given greater weight in analysis (Garson, 2007). Although the rating scales included in the current analysis share a common standardized scale (T-scores), the direct observation data were unique, reported in terms of the frequency (or rate) of various behaviors. For this reason, raw scores for each variable were used in the current analysis and then standardized as a preliminary step in data analysis. Standardization of each variable was accomplished by converting raw scores (including direct observation data) into Z-scores; the resulting Z-scores were used in all subsequent analyses.

Research Question 1: In what way(s) do dimensional measures of behavior, including the Child Behavior Checklist (Achenbach, 1991a), Teacher Report Form (Achenbach & Edelbrock, 1986), Social Skills Rating System (Gresham & Elliott, 1990), Conners' Rating Scales (Conners,

1997), and direct observation differentiate distinct subgroups within a larger group of preschoolers who meet criteria for ADHD with or without comorbid ODD? Cluster analysis was used to identify distinct subgroups within the larger sample of preschoolers who met DISC-IV criteria for ADHD and/or ODD. Hierarchical cluster analysis was selected as most appropriate for the current sample, given the relatively small sample size (< 250) and the nature of the data (interval or ratio properties; Garson, 2007). Squared Euclidean distance was selected as the most appropriate similarity measure, given its sensitivity to differences in cluster profile and score elevation (Edelbrock, 1979). In addition, squared Euclidean distance is a conventional approach, being the most common selection (Garson, 2012). Ward's method was selected as the clustering algorithm, based on its sensitivity to profile elevation (Morey, Blashfield & Skinner, 1983), which was of particular interest in the current analysis. Research has also suggested that Ward's method has particularly strong discriminatory power when combined with a Euclidean distance measure (Speece, 1994-1995).

Three criteria were used to identify the most appropriate cluster solution: the agglomeration schedule, the dendrogram, and cluster size (i.e., the number of participants within each cluster). The agglomeration schedule was examined first. When Ward's method is applied, the coefficients on the agglomeration schedule reflect the within-cluster sum of squares at each stage of the clustering process (Norušis, 2009). At each step, clustering is based on the combination of clusters/variables resulting in the smallest increase in within-cluster variability (Garson, 2012). Examination of the change in coefficient values was used to assist in identifying the most appropriate cluster solution, as larger values indicate a greater increase in the overall within-cluster sum of squares, suggesting a less optimal solution.

The dendrogram, a visual depiction of the distance between clusters at the point at which

they are merged, also facilitates selection of the most appropriate cluster solution (Garson, 2012; Norušis, 2009). The dendrogram is automatically rescaled by the Statistical Product and Service Solutions (SPSS) software program such that the distances between clusters fall between 1 and 25 (the standard range employed by SPSS). The relative distances between clusters retain their true ratios (Norušis, 2009). The vertical lines on the dendrogram indicate joined clusters. Shorter lines indicate that cases/clusters of close proximity (i.e., greater similarity) were merged; longer lines suggest the merging of more distant (and thus relatively dissimilar) cases/clusters (Garson, 2012).

Cluster size (i.e., the number of participants comprising a given cluster) is also an important consideration in selecting the most appropriate solution (Garson, 2012; Sharma & Kumar, 2006; Zhu, Wang & Li, 2010). A solution resulting in one or more very small clusters suggests that too many clusters have been requested/selected for inclusion; conversely, solutions that include a large, dominant cluster may consist of too few clusters (Garson, 2012). One heuristic that has been suggested is that the solution be comprised of clusters retaining a size ratio of approximately 1:3, such that the largest cluster is no more than three times the size of smallest cluster (Gaskin, 2012). Relative cluster size/ratio was the final criterion used to assist in selecting the most appropriate cluster solution. The nature of the obtained clusters was examined via subsequent analyses.

Research Question 2: What salient characteristics will differentiate among the identified subgroups? Chi square analyses were conducted to determine whether the participants in each cluster differed significantly with regard to relevant demographic characteristics (Table 7). Following this, a one-way multivariate analysis of variance (MANOVA) was conducted to determine the presence of significant between-cluster differences with regard to the target

measures. Cluster membership was entered as the independent (grouping) variable. The dependent variables were all of the behavioral measures included in the cluster analysis (listed in Table 8); specifically, selected subscales from several parent and teacher rating forms (Achenbach CBCL and TRF, Conners-P, Conners-T, SSRS-P, SSRS-T), as well as direct observation data gathered via the Abikoff Classroom Observation Code and the Early Screening Project observation systems.

For variables on which significant between-cluster differences occurred, the nature of the differences was examined. As noted previously, the data were converted to Z-scores prior to analysis, in order to prevent undue weighting of variables with larger standard deviations (Garson, 2007). The mean Z-score of each cluster was calculated for every variable that evidenced significant differences between clusters (Table 9). Z-scores, ranging in value from –1.0 to 1.0, have a mean of 0 and a standard deviation of 1.0. A positive mean Z-score indicated that on average, the participants in a given cluster had elevated rating scale scores or elevated levels of the observed behavior. Conversely, a negative mean Z-score reflected lower scores/ lower levels of a given behavior among the participants within the relevant cluster. The salient characteristics of each cluster were ascertained via examination of the pattern of scores that characterized each cluster on the variables for which statistically significant between-cluster occurred, and via analysis of the nature of the differences between the clusters (i.e., which variables distinguished between them).

Research Question 3: Will classifications based on cluster analysis be consistent with classifications based on DSM-IV diagnostic categories? The DBD portion of the DISC-IV was administered via parent interview (as part of the screening measures). Based on parent-reported symptomology, the participants were identified as meeting diagnostic criteria for one of four

mutually exclusive diagnostic categories: ADHD Inattentive Type, ADHD Hyperactive-Impulsive Type, ADHD Combined Type, or comorbid ADHD + ODD (as noted previously, none of the participants met criteria for a diagnosis of ODD only, per study inclusion criteria). These diagnostic classifications were compared with the cluster solution to determine the degree of overlap between groupings identified categorically (via DISC-IV diagnosis) and dimensionally-based groupings (i.e., clusters). The number of groups emerging from each classification method was considered, as was the diagnostic composition of each cluster. A quantitative comparison was also performed. Chi square analysis was conducted to determine whether there was a significant relationship between classifications based on dimensional measures (i.e., cluster membership) and categorical classifications based on the DSM-IV diagnostic categories.

Chapter IV: Results

Preliminary Data Analysis: Multicollinearity.

Statistically significant correlations among the variables initially selected for inclusion in the cluster analysis (Table 5) ranged from $r = 0.22$ ($p < 0.05$) to $r = 0.93$ ($p < 0.01$). Because strong correlations between variables can result in undue weighting of constructs that are over-represented (Sambandam, 2003), the variables with high correlations ($r \geq \left| 0.60 \right|$) were analyzed for potential redundancy. This analysis led to the removal of 30 variables. Among the remaining 43 variables retained for further analysis, statistically significant correlations ranged from $r = 0.24$ ($p = 0.048$) to $r = 0.598$ ($p < 0.001$).

Research Question 1: In what way(s) do dimensional measures of behavior differentiate distinct subgroups within a larger group of preschoolers who meet criteria for ADHD with or without comorbid ODD?

The agglomeration schedule (Table 6) reflects the within-cluster sum of squares at each stage of the clustering process. As cases/clusters are combined, the coefficients increase. Throughout most stages of the clustering process, the changes in coefficient value were relatively small and gradual, increasing by increments of 0.01 to 13.26. A greater increase in coefficient value was evident when transitioning from a three-cluster solution to a two-cluster solution (change of 31.50), suggesting a relatively larger increase in the within-cluster sum of squares. The transition from a two-cluster to a single-cluster solution was associated with a significantly larger change in coefficient value (change of 151.23), indicating that the one-cluster solution resulted in a substantial increase in within-cluster sum of squares. This signified that the one-cluster solution involved the merging of dissimilar clusters, suggesting that a two-cluster solution would be more appropriate.

In order to facilitate interpretation, the results of the agglomeration schedule were graphically displayed (Figure 1), providing a visual depiction of the change in within-cluster sum of squares corresponding with the number of clusters comprising a given solution. The graph is characterized by very short distances between data points (reflecting small changes in sum-of-square values) throughout most of the clustering stages. The distance lengthens between the three- and the two-cluster solutions, suggesting that the two-cluster solution was associated with a somewhat larger increase in within-cluster variability. Visual examination confirmed that the transition from a two-cluster solution to a single cluster was associated with a significantly larger increase in within-cluster sum of squares (as evident in a substantial lengthening in the connecting line), supporting the two-cluster solution.

The dendrogram produced by the cluster analysis (rescaled to values between 1 and 25; Figure 2) was also analyzed to assist in identifying the most appropriate cluster solution. The three-cluster solution was associated with distance values ≤ 10. The two-cluster solution was associated with distance values up to 13, indicating that clusters of relatively close proximity were merged to produce the two-cluster solution. The one-cluster solution was associated with a distance value of 25, reflecting the merging of dissimilar clusters. These results provided further support for a two-cluster solution.

Relative cluster size/ratio was the final criterion used in determination of the most appropriate cluster solution. The three-cluster solution was characterized by highly uneven cluster sizes (N = 6, N = 33, and N = 31). The largest cluster was 5.5 times larger than the smallest cluster, violating the heuristic of maintaining an approximate size ratio of 1:3 between the largest and smallest clusters (Gaskin, 2012). The two-cluster solution was characterized by clusters very similar in size (N = 39, N = 31), falling well within the suggested guidelines.

Overall, there was concordance among all three criteria used to assist in determining the most appropriate cluster solution. Analysis of the agglomeration schedule (and corresponding graph), visual examination of the dendrogram, and consideration of the relative sizes of the resulting clusters collectively supported a two-cluster solution. In order to evaluate the internal validity of the cluster solution, a second hierarchical cluster analysis was conducted, limiting the solution to two clusters. This analysis led to identical participant cluster groupings, providing further support for the initial solution. The two-cluster solution was therefore selected as most appropriate.

Research Question 2: What salient characteristics will differentiate among the identified subgroups?

Chi square analysis was conducted to assess for statistically significant differences between the clusters with regard to relevant demographic characteristics (Table 7). It is important to note that the assumption regarding the minimum expected frequency in each cell was violated for all of the analyses, with the exception of comparisons of participants' age and gender. The results did not indicate any statistically significant differences between the clusters with regard to age, gender, ethnicity, parents' level of education, medication status (whether the child was taking a psychotropic medication), or the number of days per week the child attended an educational program/activity.

The results of a one-way multivariate analysis of variance (MANOVA) confirmed the presence of significant between-cluster differences [Wilks'$\Lambda = 0.095$, $F(43, 26) = 5.77$, partial $\eta^2 = 0.91$, $p < 0.01$], providing support for the internal validity of the two-cluster solution (Marsh & Williams, 2004). Follow-up univariate analyses of variance (ANOVAs) were conducted to examine the nature of the differences between the clusters. In order to control for an inflated

Type I error due to conducting multiple analyses, the alpha level for interpreting between-cluster differences as significant was adjusted to $p \leq 0.001$. The results (Table 8) revealed that significant between-cluster differences occurred on 11 of the 43 dependent variables. In order to facilitate interpretation of the results, the mean Z-score of each cluster on each of the variables reflecting significant between-cluster differences (Table 9) was graphically displayed (Figure 3).

Cluster 1: Significant Concerns. The first cluster is comprised of 39 participants and accounts for 55.71% of the sample. This cluster had a positive mean Z score on every variable assessing problem behavior on which significant between-cluster differences occurred, across parent and teacher ratings. These results reflected more highly elevated levels of problem behavior compared with participants in Cluster 2 (Figure 3). Both parents and teachers rated the participants in Cluster 1 as exhibiting (on average) higher levels of social problems as well as oppositional, delinquent and/or problem behavior. Teacher ratings also reflected higher levels of hyperactivity and thought problems, and parent ratings reflected higher levels of anxious/shy behavior and perfectionism. The only variable on which the mean Z-score among Cluster 1 participants was negative was the Social Skills subscale of the SSRS-T, indicating that participants in Cluster 1 were rated by teachers as having weaker social skills, on average, than Cluster 2 participants. Examination of these salient characteristics led to selection of the label "*Significant Concerns*" as an appropriate descriptor for Cluster 1.

Cluster 2: Moderate Concerns. The second cluster is comprised of 31 participants and accounts for 44.29% of the sample. This cluster had a negative mean Z-score on every variable assessing problem behavior on which a significant between-cluster difference occurred, across parent and teacher ratings (Figure 3). These results reflected lower levels of problem behavior in comparison with Cluster 1 participants. Specifically, the participants in Cluster 2 were rated by

parents and teachers as exhibiting (on average) lower levels of social problems as well as oppositional, delinquent and/or problem behavior. Teacher ratings reflected lower levels of hyperactivity and thought problems, and parent ratings reflected lower levels of anxious/shy behavior and perfectionism. The sole variable on which participants in Cluster 2 had a positive mean Z-score was the Social Skills subscale of the SSRS-T, indicating that they were rated by their teachers as exhibiting stronger social skills than participants in Cluster 1. Based on these characteristics, "*Moderate Concerns*" was selected as an appropriate descriptor for Cluster 2.

Research Question 3: Will classifications based on cluster analysis be consistent with classifications based on DSM-IV diagnostic categories?

Visual comparison of the cluster solution and participants' diagnostic status indicated a lack of correspondence between the two approaches to grouping. Four diagnostic groups were represented (ADHD: Inattentive, ADHD: Hyperactive/Impulsive, ADHD: Combined, and comorbid ADHD + ODD), whereas the dimensional measures suggested only two clusters. This discrepancy indicates that factors other than ND/DBD diagnostic criteria accounted for the differences between the empirically-derived subgroups (i.e., clusters).

Chi square analysis was conducted to determine (quantitatively) whether there was a significant relationship between dimensionally-based cluster membership and categorical classifications based on the DSM-IV diagnostic categories (Table 10). It is important to note that the assumption that the expected frequency in each cell be at least five participants (McCall, 2000) was violated for many of the analyses. For both clusters, this value fell below five with regard to the expected number of participants within each of the following diagnostic groups: ADHD: Inattentive, ADHD: Hyperactive/Impulsive, and ADHD: Combined. For this reason, the results should be interpreted with caution. The chi square analysis did not reflect a significant

relationship between cluster membership and diagnostic classification for any of the diagnostic categories (ADHD-Inattentive Type, ADHD-Hyperactive-Impulsive type, ADHD-Combined Type, ODD, or comorbid ADHD + ODD). This was consistent with the initial conclusions suggested by visual analysis.

The diagnostic composition of each cluster was evaluated to allow for further examination of the relationship between diagnostic status and cluster membership. The percentage of participants within each cluster who met DISC-IV criteria for various ND/DBD diagnoses is graphically depicted in Figures 4 and 5. The diagnostic composition of both clusters was highly similar. Across both clusters, a strong majority of participants met criteria for comorbid diagnoses of both ADHD and ODD (Cluster 1: 77%, Cluster 2: 81%). Few participants in either cluster met criteria solely for the Inattentive Type of ADHD (Cluster 1: 2%, Cluster 2: 3%), and no participants met criteria for ODD-only. ADHD: Hyperactive-Impulsive Type and ADHD: Combined Type each accounted for relatively small percentages of participants in each cluster (Cluster 1: 13% and 8%, respectively; Cluster 2: 6% and 10%, respectively).

Overall, the results reflected inconsistency in the number of subgroups suggested by a categorical versus dimensional approach, the absence of a statistically significant relationship between cluster membership and diagnostic status, and a high level of similarity in the diagnostic composition of each cluster. It was also noted that the (dimensional) characteristics distinguishing between the two clusters extended beyond the traits typically associated with ADHD and/or ODD. Compared with participants in Cluster 2 (*Moderate Concerns*), the participants in Cluster 1 (*Significant Concerns*) exhibited elevated levels hyperactivity (per teacher rating), and oppositional/defiant and problem behavior (per parent and teacher ratings). However, they also exhibited (on average) more social problems (per parent and teacher ratings)

and elevated levels of anxious/shy behavior and perfectionism (per parent ratings). In contrast, Cluster 2 participants exhibited lower levels of problem behavior across the areas assessed and were rated by teachers as exhibiting stronger social skills.

Chapter V: Discussion

A growing literature base indicates that problem behaviors emerge at an early age, often to a sufficient degree of severity to warrant formal diagnosis (American Academy of Pediatrics, 2011). The rates of psychiatric disorders among preschoolers have been found to be consistent with those reported among school-aged children (Egger & Angold, 2006). Neurodevelopmental and disruptive behavior disorders (particularly ADHD and ODD, respectively) are among the most common diagnoses within the preschool population. Rather than reflecting a developmental phase, evidence indicates that early-onset behavior problems often persist over time (Campbell et al., 1984; Keenan et al., 2011; Pierce et al., 1999) and are associated with subsequent impairment (Speltz et al., 1999). Although the importance of early intervention is widely acknowledged, there are several obstacles to effective diagnosis at the preschool level, including the need to differentiate significant behavior problems from developmentally typical behavior, variations in behavior across settings, the difficulty of gaining information from multiple sources, and questions regarding the relative utility of categorical and dimensional approaches to assessment and diagnosis. The purpose of this study was to examine the ways in which dimensional measures of behavior (parent ratings, teacher ratings, and direct observation) differentiated among preschoolers who met DISC-IV criteria for ADHD, with or without comorbid ODD.

Research Question 1 & Hypothesis: Dimensionally-Based Differentiation among Subgroups

The first research question focused on the ways in which dimensional measures of behavior (including rating scales and direct observation data) differentiated distinct subgroups within a sample of preschoolers who met diagnostic criteria for ADHD, with or without comorbid ODD. It was hypothesized that cluster analysis would identify a greater number of homogeneous subgroups among the participants than would directly correspond with the diagnostic categories

of ODD and the various subtypes of ADHD, as defined in the DSM-IV. This hypothesis was not supported; hierarchical cluster analysis indicated two clusters, fewer than would correspond with the DSM-IV diagnostic categories represented within the sample. It is important to note that the relatively small size and homogeneous diagnostic composition of the sample likely impacted the results. A strong majority of the participants met DISC-IV criteria for comorbid ADHD and ODD (N = 55, 78.6%). The relatively homogeneous nature of the sample with regard to diagnostic status likely contributed to the limited cluster groupings.

In direct contrast with a categorical approach to diagnosis in which symptoms are identified as either present or absent, it was hypothesized that the relative severity of behavioral and/or emotional characteristics would contribute to the distinction among subgroups. This hypothesis was supported. A MANOVA conducted to examine the differences between the two clusters revealed significant between-cluster differences 11 of the 43 variables included in analysis. The nature of the differences between the two clusters was highly consistent across variables. The first cluster was characterized by more elevated average scores (reflecting higher levels of concern) on each of the variables assessing problem behavior on which a significant between-cluster difference occurred, across parent and teacher ratings. Based on these results, *Significant Concerns* was selected as an appropriate label for Cluster 1, as participants exhibited more problem behaviors than typical peers (per inclusion criteria) and higher levels of problem behavior than Cluster 2 participants (per parent and teacher ratings).

The second cluster was comprised of participants rated as exhibiting higher levels of inattention and problem behaviors than typical peers (per inclusion criteria). However, Cluster 2 participants were rated as exhibiting lower levels of problem behavior than those in Cluster 1, as evidenced by a lower average score on each measure of problem behavior on which the clusters

differed significantly. This led to selection of *Moderate Concerns* as an appropriate label for Cluster 2. One variable on which significant between-cluster differences occurred was relatively unique, as it assessed skill level, rather than evaluating a problem behavior. Teachers rated the participants in Cluster 2 as exhibiting stronger social skills (on average) than participants in Cluster 1 (per the *SSRS-T*). Parent ratings of social skills (via the *SSRS-P*) did not differ significantly between the clusters.

Overall, these results are consistent with previous research findings. The majority of the differences between the two clusters pertained to the severity of the problems reported, rather than their presence or absence. This finding coincides with the widely touted concept that because many of the symptoms of ADHD and DBD are developmentally typical for preschoolers (Byrne et al., 2000; Connor, 2002; Keenan & Wakschlag, 2000; Keenan et al., 1997), such factors as severity, frequency, and persistence are important factors to consider in assessment and diagnosis (Carter et al., 2004; Kadesjo et al., 2003).

Research Question 2 & Hypotheses: Relationship between Parent and Teacher Ratings

Based on prior evidence of inconsistency between parent and teacher ratings (Achenbach et al., 2008; Drabick et al., 2004; Jensen et al., 1993; Kerr, Lukenheimer & Olson, 2007; Serna et al., 2002), it was hypothesized that the parent and teacher ratings characterizing the various clusters would differ. In general, this hypothesis was not supported. Across nearly every variable, parent and teacher ratings followed a consistent pattern, reflecting (on average) more elevated concerns for participants in Cluster 1 than Cluster 2. Among the 11 variables evidencing statistically significant between-cluster differences, two comprised a corresponding pair of parent- and teacher- rated subscales (SSRS-P: Problem Behaviors and SSRS-T: Problem Behaviors). For seven of the variables, corresponding parent and teacher ratings trended in the

same direction, with more elevated scores for Cluster 1 than for Cluster 2 participants (Table 11).

For the remaining two variables, either the parent rating subscale or the corresponding teacher rating subscale was removed from analysis to prevent multicollinearity, precluding comparison of potential correspondence between parent and teacher ratings (Table 11). Specifically, Cluster 1 participants were rated by parents as exhibiting higher levels of oppositional behavior than Cluster 2 participants (per the *Conners-P*), and were rated by teachers as exhibiting more thought problems (per the *TRF*). For both variables, the corresponding teacher- or parent-rating subscale was removed prior to analysis (specifically, *Conners-T*: Oppositional and *CBCL*: Thought Problems). For this reason, the potential similarity between parent and teacher ratings could not be ascertained.

Overall, there was a high level of similarity between parent and teacher ratings. However, it is important to note that this may have been an artifact of inclusion criteria, which required elevated scores (1.5 standard deviation above the mean) on both parent and teacher ratings via the Conners Rating Scales, Revised. As a result, all participants were perceived by both parents and teachers as exhibiting symptoms of ADHD to a greater degree than typical peers, increasing the likelihood that similarities would be evident on parent and teacher ratings of other behaviors.

Research Question 3 & Hypotheses: Relationship between Clusters and DSM-IV Diagnostic Categories

In light of previous findings of a relationship between rating scale measures and various DSM diagnoses (Biederman et al., 1993; Chen et al., 1994; Edelbrock & Costello, 1988; Frankel et al., 1992; Hudziak et al., 2004), it was hypothesized that the clusters would share some similarities with DSM-IV diagnostic criteria (e.g., patterns of elevation on scales and measures corresponding with the characteristics included in the DSM-IV) and that participants with the

same diagnosis(es) would tend to be grouped within the same cluster. These hypotheses were not supported. Chi square analysis did not reflect a significant relationship between cluster membership and diagnostic classification for any of the diagnostic categories. As noted previously, the primary differences between the two clusters pertained to degree of score elevation, rather than distinct patterns of behaviors or traits unique to a given cluster. In addition, the diagnostic composition of the two clusters was highly similar. Both were comprised primarily of participants who met DISC-IV criteria for comorbid ADHD and ODD. This is consistent with previous findings that among preschoolers with an ND/DBD, comorbidity is more common than having a single diagnosis (Kaplan et al., 2001; Wilens et al., 2002) and that ADHD and ODD is the most common dual diagnosis (Kadesjo et al., 2003; Keenan & Wakschlag, 2000).

Summary of Findings

Overall, the findings indicated that the two emergent clusters were distinguished primarily by degree of problem behavior, rather than the nature of the concerns. Parent and teacher ratings followed a consistent pattern, with participants in Cluster 1 (*Significant Concerns)* rated as exhibiting more elevated problems than those in Cluster 2 (*Moderate Concerns*) with regard to externalizing problems (delinquent/oppositional behavior, problem behavior) and social problems. Parent ratings of anxious/shy behavior and perfectionism also reflected a higher level of concern among participants in Cluster 1 than those in Cluster 2. Corresponding teacher ratings followed the same trend, but did not reach statistical significance. Similarly, teacher ratings indicated that participants in Cluster 1 exhibited, on average, higher levels of hyperactivity than those in Cluster 2, with parent ratings of hyperactivity reflecting the same trend, but failing to reach statistical significance.

One relatively unique finding was that participants in Cluster 2 were rated by teachers (on the SSRS-T) as exhibiting more positive social skills, on average, than Cluster 1 participants. [Parent ratings of social skills were not significantly different between the clusters]. Previous research has indicated that children with ADHD exhibit social skill impairments (Lahey et al., 1998). Research has also suggested a possible relationship between comorbidity and social skills impairment. A meta-analysis conducted by Waschbusch (2002) found that individuals with symptoms of both ADHD and conduct problems exhibited higher levels of aggression, longer maintenance of aggression over time, more covert antisocial behaviors, and more peer-related problems than individuals with symptoms of ADHD only. Although both clusters were comprised primarily of children meeting DISC-IV diagnostic criteria for both ADHD and ODD, given evidence of fluctuations in ADHD symptomology and subtype classification over time (Lahey et al., 2005; Lahey & Willcutt, 2010), it is possible that the diagnostic status of the current participants could change over time. Current between-cluster differences may relate to later-emerging differences in diagnostic status. As comorbidity is associated with increased number and severity of symptoms among preschoolers (Cunningham & Boyle, 2002; Gadow & Nolan, 2002), participants in Cluster 1 may be at increased risk for maintaining a dual diagnosis of ADHD and ODD than those grouped in Cluster 2.

Implications

Consistent with previous research, the current results indicate that symptoms of ADHD and ODD emerge at an early age, and often co-occur. The majority of the participants (78.57%) met criteria for comorbid ADHD and ODD. Although this may have influenced by participant inclusion criteria (further discussed below), this is consistent with research indicating that a high percentage of preschoolers with ADHD (up to 60-70%) also meet criteria for comorbid ODD

(Kadesjo et al., 2003; Keenan & Wakschlag, 2000). The findings suggest that preschoolers who display elevated symptoms in one area (e.g., ADHD characteristics) are likely to exhibit problems in other symptom domains, particularly if concerns are evident across settings.

The current findings support the conclusion that assessment should incorporate evaluation of a wide range of behaviors and characteristics. Although recruited for symptoms of ADHD, the participants exhibited symptoms across a variety of domains, including oppositional and defiant behavior, social difficulties, and internalizing problems (e.g., anxious/shy behavior, perfectionism, and thought problems). The results suggest that young children presenting with elevated levels of inattention and/or hyperactivity may also display problems in other areas, warranting a comprehensive approach to assessment. In addition to gathering information about problem behaviors and other concerns across a variety of domains, skills assessment also appears to be an important component of evaluation (Carter et al., 2004). The current findings indicated that teacher ratings of social skills differed significantly between the two clusters, suggesting that this skill domain may have important implications. Variations in skill acquisition and application may serve as an early indicator of increased risk for adverse outcomes. A comprehensive approach to assessment is important not only for practitioners completing evaluations for diagnostic purposes, but also for professionals working with young children in a range of capacities. For school-based practitioners, it is important to note that students presenting with symptoms in one domain are likely struggling in other areas as well. Evaluation should address multiple symptom domains, including skills assessment.

These results also have implications for treatment, particularly from a multi-systems perspective. In light of the findings that the participants displayed elevated symptoms in multiple areas, it is important that treatment address needs across domains. Interventions for children who

exhibit characteristics of ADHD would likely be most effective if they also address behavioral concerns (e.g., targeting oppositional/defiant behavior) as well as social skills (teaching social skills and strengthening peer relationships). Intervention for internalizing problems may also be warranted. The current results also suggest the importance of providing support across settings. Elevated symptoms were reported by both parents and teachers, underscoring the importance of a comprehensive approach to treatment that provides support both at home and in school.

Limitations

The results of this study should be interpreted within the context of several limitations. One notable limitation is the small sample size. It is possible that data from a larger sample would suggest a greater number of clusters differentiated not only by symptom severity, but possibly also by the nature of the concerns prominent within each cluster. Because the small sample size resulted in lower-than-optimal expected cell counts, the results of the chi square analysis must be interpreted with caution.

Factors related to participant selection may also have impacted the current results. Because the purpose of longitudinal study (of which these data are a subset) was to evaluate treatment outcomes for children with ADHD, inclusion criteria required elevated scores on both parent and teacher ratings via the Conners Rating Scales, Revised. This contributed to the similarity between parent and teacher ratings and resulted in a restricted range of scores among participants. In addition, it is possible that screening for an elevated level of ADHD symptoms increased the likelihood the participants would also meet criteria for a comorbid disorder (specifically, ODD). The relatively stringent nature of the inclusion criteria may have impacted the current study results and also poses a limitation to the external validity of the results, which may apply only to children with elevated symptoms across settings and/or who exhibit symptoms

of both ADHD and ODD.

It is also important to note that the current sample was relatively homogenous with regard to gender, ethnicity, and diagnostic status. The homogeneity of the sample's diagnostic composition limited the power of the current study to examine the relationship between diagnosis and cluster membership. Further research is needed to determine whether similar results would be obtained within a larger and more diverse sample. The homogeneity of the sample also poses a limitation to the external validity of the current results. As noted above, participants were recruited based on meeting DISC-IV diagnostic criteria for ADHD. For this reason, although many of the participants met criteria for both ADHD and ODD, none met criteria for ODD alone. Similarly, children with Autism and Conduct Disorder were excluded from the study, precluding determination of whether similar findings would be evident within these populations.

Certain measures used in the current study also pose a limitation, as they do not reflect best (or typical) practice for a clinical evaluation conducted for diagnostic purposes. Due to their selection for use in a longitudinal study, several of the measures were normed on a population older than the participants were at the time baseline data (used in the current study) were gathered. The Achenbach CBCL was normed on children ages 4- 18 years, and the TRF rating scale was normed on children ages 5 – 18 years; thus, the rating scales were age-appropriate for some, but not all, of the participants. This was not the case for the SSRS, given that the Elementary Level was administered. Comparison of the Preschool and Elementary Levels of the SSRS revealed discrepancies in item content. Perhaps more significantly, research has suggested that the SSRS might not assess the same construct(s) over time (Van Horn, Atkins-Burnett, Karlin, Ramey & Snyder, 2007). These issues raise questions regarding whether the rating scale assessed the intended traits among the current sample participants.

Future Directions

The generally high level of agreement between parent and teacher ratings was unexpected, in light of previous research findings (Achenbach et al., 2008; Drabick et al., 2004; Jensen et al., 1993; Serna et al., 2002). This outcome could be an artifact of inclusion criteria, but might also be related to the young age of the participants. With time, the school experience becomes increasingly dissimilar from the home setting, with progressively increased academic demands and heightened expectations for independence and self-regulation. As these changes occur, differences between parent and teacher ratings will likely become more pronounced. Future research might address whether the degree of correspondence between parent and teacher ratings is related to the age of the child being assessed.

In light of the limitations described above, it would be beneficial for the current study to be replicated with a larger, more heterogeneous sample. More equivalent sample representation across diagnostic categories would allow for a more comprehensive examination of the relationship between dimensional measures and diagnostically-based categorical groupings. The inclusion of a control comparison group might help clarify the ways in which dimensional measures differentiate among subgroups (e.g., presence/absence of diagnosis, symptom presence/severity, comorbidity).

As noted previously, the current study analyzed data gathered at baseline as part of a larger, longitudinal treatment study. Future research could examine the relationship between early cluster membership (at baseline) and subsequent outcomes, including diagnostic status, as manifested over time. In addition to assessing ND and DBD symptom manifestation and severity, it would likely be beneficial for subsequent research to include assessment of internalizing symptoms, measures of impairment, as well as skills assessment, thus providing for

a more comprehensive analysis of participants' functioning across important domains.

Subsequent research might also investigate potential between-cluster differences in treatment response. Such findings would have implications for early identification, helping to ascertain the domains that are important to address during assessment as well as the sources of information and methods of data collection that might yield the most useful information. Ultimately, such information will contribute to more efficient and accurate early identification, facilitating selection of the most appropriate form(s) of early intervention.

References

Abikoff, H. & Gittelman, R. (1985). The normalizing effects of methylphenidate on the classroom behavior of addh children. *Journal of Abnormal Child Psychology, 13(1),* 33-44.

Abikoff, H., Gittelman-Klein, R. & Klein, D.F. (1977). Validation of a classroom observation code for hyperactive children. *Journal of Consulting and Clinical Psychology, 45(5),* 772-783.

Abikoff, H., Gittelman, R. & Klein, D.F. (1980). Classroom observation code for hyperactive children: A replication of validity. *Journal of Consulting and Clinical Psychology, 48(5),* 555-565.

Abikoff, H.B., Jensen, P.S., Arnold, L.L.E., Hoza, B., Hechtman, L., Pollack, S., Martin, D., Alvir, J., March, J.S., Hinshaw, S., Vitiello, B., Newcorn, J., Greiner, A., Cantwell, D.P., Conners, C.K., Elliott, G., Greenhill, L.L., Kraemer, H., Pelham, W.E., Severe, J.B., Swanson, J.M., Wells, K. & Wigal, T. (2002). Observed classroom behavior of children with adhd: Relationship to gender and comorbidity. *Journal of Abnormal Child Psychology, 30(4),* 349-359.

Achenbach, T.M. (1991a). *Integrative guide for the 1991 cbcl/4-18, ysr, and trf profiles.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T.M. (1991b). *Manual for the child behavior checklist/4-18 and 1991 profile.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M. (2001). Challenges and benefits of assessment, diagnosis, and taxonomy for clinical practice and research. *Australian and New Zealand Journal of Psychiatry, 35(3),* 263-271.

Achenbach, T.M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H-C. &

Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with aseba and sdq instruments: Research findings, applications, and future directions. *Journal of Child and Adolescent Psychiatry, 49(3),* 251-275.

Achenbach, T.M., Dumenci, L. & Rescorla, L.A. (2003). DSM-oriented and empirically-based approaches to constructing scales from the same item pools. *Journal of Clinical Child and Adolescent Psychology, 32(3),* 328-340.

Achenbach, T. M. & Edelbrock, C. (1986). *Manual for the teacher report form.* Burlington, Vermont: University of Vermont, Department of Psychiatry. As cited in Campbell & Ewing, 1990.

Achenbach, T.M. & Edelbrock, C. (1987). *Manual for the youth self-report and profile.* Burlington: University of Vermont, Department of Psychiatry. As cited in Pierce et al., 1999.

American Academy of Pediatrics (2011). ADHD: Clinical practice guidelines for the diagnosis, evaluation, and treatment of attention deficit/hyperactivity disorder in children and adolescents. *Pediatrics, 128 (5),* 1-16.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text rev.).* Washington, DC: Author.

Arons, B.S., Katz-Leavy, J., Wittig, A.C. & Holden, E.W. (2002). Too young for adhd: The potential role of systems of care (attention-deficit hyperactivity disorder). *Journal of Developmental and Behavioral Pediatrics, 23(1),* 57-63.

Banaschewski, T. (2010). Editorial: preschool behavior problems – over-pathologised or under-identified? A developmental psychopathology perspective is needed. *Journal of Child Psychology and Psychiatry, 51(1),* 1-2.

Behar, L. (1977). The preschool behavior questionnaire. *Journal of Abnormal child Psychology,*

5, 83-96.

Biederman, J., Faraone, S.V., Doyle, A., Lehman, B.K., Kraus, I., Perrin, J. & Tsuang, M.T. (1993). Convergence of the child behavior checklist with structured interview-based psychiatric diagnoses of adhd children with and without comorbidity. *Journal of Child Psychology and Psychiatry, 34(7),* 1241-1251.

Biederman, J., Spencer, T.J., Newcorn, J.H., Gao, H., Milton, D.R., Feldman, P.D. & Witte, M.M. (2006). Effect of comorbid symptoms of oppositional defiant disorder on responses to atomoxetine in children with adhd: A meta-analysis of controlled clinical trial data. *Psychopharmacology, 190(1),* 31-41.

Brandau, H. & Pretis, M. (2004). Early identification and systemic educational intervention for young children with attention-deficit/hyperactivity disorder (ad/hd). *European Journal of Special Needs Education, 19(1),* 17-31.

Burns, G.L., Walsh, J.A., Owen, S.M. & Snell, J. (1997). Internal validity of attention deficit hyperactivity disorder, oppositional defiant disorder, and overt conduct disorder symptoms in young children: Implications from teacher ratings for a dimensional approach to symptom validity. *Journal of Clinical Child Psychology, 26(3),* 266-275.

Burns, G.L. & Walsh, J.A. (2002). The influence of adhd-hyperactivity/impulsivity symptoms on the development of oppositional defiant disorder symptoms in a 2-year longitudinal study. *Journal of Abnormal Child Psychology, 30(3),* 245-256.

Bussing, R., Lehninger, F. & Eyberg, S. (2006). Difficult child temperament and attention-deficit/hyperactivity disorder in preschool children. *Infants & Young Children, 19(2),* 123-131.

Byrne, J.M., Bawden, H.N., Beattie, T.L. & DeWolfe, N.A. (2000). Preschoolers classified as

having attention-deficit hyperactivity disorder (adhd): dsm-iv symptom endorsement pattern. *Journal of Child Neurology, 15(8),* 533-538.

Campbell, S.B., Breaux, A.M., Ewing, L.J. & Szumowski, E.K. (1984). A one-year follow-up study of parent-referred hyperactive preschool children. *Journal of the American Academy of Child Psychiatry, 23(3),* 243-249.

Campbell, S.B., Breaux, A.M., Ewing, L.J. & Szumowski, E.K. (1986a). Correlates and predictors of hyperactivity and aggression: A longitudinal study of parent-referred problem preschoolers. *Journal of Abnormal Child Psychology, 14(2),* 217-234.

Campbell, S.B., Ewing, L.J., Breaux, A.M. & Szumowski, E.K. (1986b). Parent-referred problem three-year-olds: Follow-up at school entry. *Journal of Child Psychology and Psychiatry 27(4),* 473-488.

Campbell, S.B. & Ewing, L.J. (1990). Follow-up of hard-to-manage preschoolers: Adjustment at age 9 and predictors of continuing symptoms. *Journal of Child Psychology and Psychiatry 31(6),* 871-889.

Carter, A.S., Briggs-Gowan, M.J. & Davis, N.O. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry, 45(1),* 109-134.

Chen, W.J., Faraone, S.V., Biederman, J. & Tsuang, M.T. (1994). Diagnostic accuracy of the child behavior checklist scales for attention-deficit hyperactivity disorder: A receiver-operating characteristic analysis. *Journal of Consulting and Clinical Psychology, 62(5),* 1017-1025.

Cluett, S.E., Forness, S.R., Ramey, S.L., Ramey, C.T., Hsu, C., Kavale, K.A. & Gresham, F.M. (1998). Consequences of differential diagnostic criteria on identification rates of

children with emotional or behavior disorders. *Journal of Emotional and Behavioral Disorders, 6(3),*130-140.

Conners, C.K. (1997). *Conners' rating scales – revised: Technical manual.* North Tonawanda, NY: Multi-Health Systems.

Conners, C.K., Sitarenios, G., Parker, J.D.A. & Epstein, J.N. (1998a). The revised conners' parent rating scale (cprs-r): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology, 26(4),* 257-268.

Conners, C.K., Sitarenios, G., Parker, J.D.A. & Epstein, J.N. (1998b). Revision and restandardization of the conners' teacher rating scale (ctrs-r): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology, 26(4),* 279-291.

Connor, D.F. (2002). Preschool attention deficit hyperactivity disorder: A review of prevalence, diagnosis, neurobiology, and stimulant treatment. *Journal of Developmental and Behavioral Pediatrics, 23(1),* 1-10.

Costello, A.J., Edelbrock, C., Kalas, R., Kessler, M.D. & Klaric, S.H. (1982). *The nimh diagnostic interview schedule for children (disc). Unpublished interview schedule.* Department of Psychiatry, University of Pittsburgh. As cited in Campbell & Ewing, 1990.

Cunningham, C.E. & Boyle, M.H. (2002). Preschoolers at risk for attention deficit hyperactivity disorder and oppositional defiant disorder: Family, parenting, and behavioral correlates. *Journal of Abnormal Child Psychology, 30(6),* 555-569.

Drabick, D.A.G., Gadow, K.D., Carlson, G.A. & Bromet, E. J. (2004). ODD and adhd symptoms in ukranian children: External validators and comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry, 43(6),* 735-743.

Edelbrock, C. (1979). Mixture model tests of hierarchical clustering algorithms: The problem of

classifying everybody. *Multivariate Behavioral Research, 14,* 367-384.

Edelbrock, C. & Costello, A.J. (1988). Convergence between statistically derived behavior problem syndromes and child psychiatric diagnoses. *Journal of Abnormal Child Psychology, 16(2),* 219-231.

Egger, H.L. & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry, 47(3-4),* 313-337.

Feil, E.G. & Becker, W.C. (1993). Investigation of a multiple-gated screening system for preschool behavior problems. *Behavioral Disorders, 19(1),* 44-53.

Feil, E.G., Severson, H.H. & Walker, H.M. (1998). Screening for emotional and behavioral delays: The early screening project. *Journal of Early Intervention, 21(3),* 252-266.

Fergusson, D.M. & Horwood, J. (1995). Predictive validity of categorically and dimensionally scored measures of disruptive childhood behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry, 34(4),* 477-485.

Frankel, F., Hanna, G.L., Cantwell, D.P., Shekim, W. & Ornitz, E. (1992). Cluster analysis of child behavior checklists of 6 to 11-year-old males with varying degrees of behavior disorders. *Child Psychiatry and Human Development, 23(2),* 69-85.

Gadow, K.D. & Nolan, E.E. (2002). Differences between preschool children with odd, adhd, and odd+adhd symptoms. *Journal of Child Psychology and Psychiatry, 43(2),* 191-201.

Gadow, K.D., Sprafkin J. & Nolan, E.E. (2001). DSM-IV symptoms in-community and clinic preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 40(12),* 1383-1392.

Garson, G.D. (2007). Cluster analysis. *Statnotes*. Retrieved April 29, 2007. From

http://www2.chass.ncsu.edu/ garson/PA765/clustecr.htm.

Garson, G.D. (2012). *Cluster analysis.* Asheboro, NC: Statistical Associates Publishers.

Gaskin, J. (2012). Two-step cluster analysis in spss. *Gaskination's Statistics*. Retrieved August 12, 2013. From http://youtube.com/Gaskination.

Gimpel, G. A. & Kuhn, B. R. (2000). Maternal report of attention deficit hyperactivity disorder symptoms in preschool children. *Child: Care, Health and Development, 26(3),* 163-176.

Gresham, F.M. & Elliott, S.N. (1990). *Social skills rating system: Manual.* Circle Pines, MN: American Guidance Service.

Harvey, E.A., Youngwirth, S.D., Thakar, D.A. & Errazuriz, P.A. (2009). Predicting attention-deficit/hyperactivity disorder and oppositional defiant disorder from preschool diagnostic assessments. *Journal of Counseling and Clinical Psychology, 77(2),* 349-354.

Healey, D.M., Marks, D.J. & Halperin, J.M. (2011). Examining the interplay among negative emotionality, cognitive functioning, and attention deficit/hyperactivity disorder symptom severity. *Journal of the International Neuropsychological Society, 17,* 502-510.

Hodges, K., Saunders, W.B., Kashani, J., Hamlett, K. & Thompson, R.J. (1990). Internal consistency of dsm-iii diagnoses using the symptom scales of the child assessment schedule. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 635-641. As cited in Pierce et al., 1999.

Hudziak, J.J., Copeland, W., Stanger, C. & Wadsworth, M. (2004). Screening for dsm-iv externalizing disorders with the child behavior checklist: A receiver-operating characteristic analysis. *Journal of Child Psychology and Psychiatry, 45(7),* 1299-1307.

Jensen, P.S., Salzberg, A.D., Richters, J.E. & Watanabe, H.K. (1993). Scales, diagnoses, and child psychopathology: I. CBCL and disc relationships. *Journal of the American Academy of*

Child & Adolescent Psychiatry, 32(2), 397-407.

Jensen, P.S., Watanabe, H.K., Richters, J.E., Roper, M., Hibbs, E.D., Salzberg, A.D. & Liu, S. (1996). Scales, diagnoses, and child psychopathology: II. Comparing the cbcl and the disc against external validators. *Journal of Abnormal Child Psychology, 24(2),* 151-168.

Jensen, P.S., Hinshaw, S.P., Kraemer, H.C., Lenora, N., Newcorn, J.H., Abikoff, H.B., March, J.S., Arnold, L.E., Cantwell, D.P., Conners, C.K., Elliott, G.R., Greenhill, L.L., Hechtman, L., Hoza, B., Pelham, W., Severe, J.B., Swanson, J.M., Wells, K.C., Wigal, T. & Vitiello, B. (2001). ADHD comorbidity findings from the mta study: Comparing comorbid subgroups. *Journal of the American Academy of Child & Adolescent Psychiatry, 40(2),* 147-158. As cited in Campbell et al., 1986.

Kadesjo, C., Hagglof, B., Kadesjo, B. & Gillberg, C. (2003). Attention-deficit-hyperactivity disorder with and without oppositional defiant disorder in 3- to 7-year-old children. *Developmental Medicine and Child Neurology, 45(10),* 693-699.

Kaplan, B.J., Dewey, D.M., Crawford, S.G. & Wilson, B.N. (2001). The term *comorbidity* is of questionable value in reference to developmental disorders: Data and theory. *Journal of Learning Disabilities, 34(6),* 555-565.

Kasius, M.C., Ferdinand, R.F., van den Berg, H. & Verhulst, F.C. (1997). Associations between different diagnostic approaches for child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 38(6),* 625-632.

Keenan, K., Boeldt, D., Chen, D., Coyne, C., Donald, R., Duax, J., Hart, K., Perrott, J., Strickland, J., Danis, B., Hill, C., Davis, S., Kampani, S., Humphries, M. (2011). Predictive validity of dsm-iv oppositional defiant and conduct disorders in clinically referred preschoolers. *Journal of Child Psychology and Psychiatry, 52 (1),* 47-55.

Keenan, K., Shaw, D.S., Walsh, B., Delliquadri, E. & Giovannelli, J. (1997). DSM-III-R disorders in preschool children from low-income families. *Journal of the American Academy of Child & Adolescent Psychiatry, 36(5),* 620-627.

Keenan, K. & Wakschlag, L.S. (2000). More than the terrible twos: The nature and severity of behavior problems in clinic-referred preschool children. *Journal of Abnormal Child Psychology, 28(1),* 33-46.

Keenan, K. & Wakschlag, L.S. (2002). Can a valid diagnosis of disruptive behavior disorder be made in preschool children? *American Journal of Psychiatry, 159(3),* 351-358.

Kern, L., DuPaul, G.J., Volpe, R.J., Sokol, N.G., Lutz, J.G., Arbolino, L.A., Pipan, M. & VanBrakle, J.D. (2007). Multisetting assessment-based intervention for young children at risk for attention deficit hyperactivity disorder: Initial effects on academic and behavioral functioning. *School Psychology Review, 36(2),* 237-255.

Kerr, D.C.R., Lukenheimer, E.S. & Olson, S.L. (2007). Assessment of child problem behaviors by multiple informants: A longitudinal study from preschool to school entry. *Journal of Child Psychology and Psychiatry, 48 (10),* 967-975.

Korkman, M., Kirk, U. & Kemp, S.L. (1998). *NEPSY – a developmental neuropsychological assessment manual.* San Antonio, TX: Psychological Corporation.

Kuhne, M., Schachar, R., Tannock, R. (1997). Impact of comorbid oppositional or conduct problems on attention-deficit hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 36(12),* 1715-1725.

Lahey, B.B., Pelham, W.E., Loney, J., Kipp, H., Ehrhardt, A., Lee, S.S., Willcutt, E.G., Hartung, C.M., Chronis, A. & Massetti, G. (2004). Three-year predictive validity of dsm-iv attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of*

Psychiatry, 161(11), 2014-2020.

Lahey, B.B., Pelham, W.E., Loney, J., Lee, S.S. & Willcutt, E. (2005). Instability of the dsm-iv subtypes of adhd from preschool through elementary school. *Archives of General Psychiatry, 62,* 896-902.

Lahey, B.B., Pelham, W.E., Stein, M.A., Loney, J., Trapani, C., Nugent, K., Kipp, H., Schmidt, E., Lee, S., Cale, M., Gold, E., Hartung, C.M., Willcutt, E. & Baumann, B. (1998). Validity of dsm-iv attention-deficit/hyperactivity disorder for younger children. *Journal of the American Academy of Child & Adolescent Psychiatry, 37(7),* 695-702.

Lahey, B.B. & Willcutt, E.G. (2010). Predictive validity of a continuous alternative to nominal subtypes of attention-deficit hyperactivity disorder for dsm-v. *Journal of Clinical Child and Adolescent Psychology, 39(6),* 761-775.

Lavigne, J.V., Gibbons, R.D., Christoffel, K.K., Arend, R., Rosenbaum, D., Binns, H., Dawson, N., Sobel, H. & Isaacs, C. (1996). Prevalence rates and correlates of psychiatric disorders among preschool children. *Journal of the American Academy of Child & Adolescent Psychiatry, 35(2),* 204-214.

Lee, S.S., Lahey, B.B., Owens, E.B. & Hinshaw, S.P. (2008). Few preschool boys and girls with adhd are well-adjusted during adolescence. *Journal of Abnormal Child Psychology, 36,* 373-383.

Mahone, E.M. & Schneider, H.E. (2012). Assessment of attention in preschoolers. *Neuropsychology Review, 22,* 361-383.

Marcus, D.K. & Barry, T.D. (2011). Does attention-deficit/hyperactivity disorder have a dimensional latent structure? A taxometric analysis. *Journal of Abnormal Psychology, 120(2),* 427-442.

Marsh, P.J. & Williams, L.M. (2004). An investigation of individual typologies of attention-deficit hyperactivity disorder using cluster analysis of dsm-iv criteria. *Personality and Individual Differences, 36,* 1187-1195.

McCall, R.B. (2000). *Fundamental statistics for behavioral sciences (8th ed.).* Belmont, CA: Wadsworth Publishing.

McGoey, K.E., DuPaul, G.J., Eckert, T.L., Volpe, R.J. & Van Brakle, J. (2005). Outcomes of a multi-component intervention for preschool children at-risk for attention-deficit/hyperactivity disorder. *Child & Family Behavior Therapy, 27(1),* 33-56.

Morey, L.C., Blashfield, R.K. & Skinner, H.A. (1983). A comparison of cluster analysis techniques within a sequential validation framework. *Multivariate Behavioral Research, 18,* 309-329.

Newcorn, J.H., Halperin, J.M., Jensen, P.S., Abikoff, H.B., Arnold, L.E., Cantwell, D.P., Conners, C.K., Elliott, G.R., Epstein, J.N., Greenhill, L.L., Hechtman, L., Hinshaw, S.P., Hoza, B., Kraemer, H.C., Pelham, W.E., Severe, J.B., Swanson, J.M., Wells, K.C., Wigal, T. & Vitiello, B. (2001). Symptom profiles in children with adhd: Effects of comorbidity and gender. *Journal of the American Academy of Child and Adolescent Psychiatry, 40(2),* 137-146.

Norušis, M.J. (2009). *SPSS statistics 17.0 statistical procedures companion.* Upper Saddle River, NJ: Prentice Hall.

Pelham, W., Gnagy, E.M., Greenslade, K.E. & Milich, R. (1992). Teacher ratings of dsm-iii-r symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 210-218. As cited in Lahey et al., 1998.

Perrin, E.C. (2000). Maternal report of attention deficit hyperactivity disorder symptoms in

preschool children: Commentary. *Child: Care, Health and Development, 26(3),* 176-178.

Pierce, E.W., Ewing, L.J. & Campbell, S.B. (1999). Diagnostic status and symptomatic behavior of hard-to-manage preschool children in middle childhood and early adolescence. *Journal of Clinical Child Psychology, 28(1),* 44-57.

Postert, C., Averbeck-Holocher, M., Beyer, T., Müller, J. & Furniss, T. (2009). Five symptoms of psychiatric classification for preschool children: Do differences in validity, usefulness and reliability make for competitive or complimentary constellations? *Child Psychiatry & Human Development, 40(1),* 25-41.

Ryan-Krause, P. (2010). Attention deficit hyperactivity disorder: Part ii. *Journal of Pediatric Health Care, 24(5),* 338-342.

Sambandam, R. (2003). Cluster analysis gets complicated. *Marketing Research, 15(1),* 16-21.

Sarason, I.G., Johnson, J.A. & Siegel, J.M. (1978). Assessing the impact of life changes: Development of the life experiences survey. *Journal of Consulting and Clinical Psychology, 46,* 932-946. As cited in Campbell & Ewing, 1990.

Serna, L., Nielsen, E., Mattern, N. & Forness, S.R. (2002). Use of different measures to identify preschoolers at-risk for emotional or behavioral disorders: Impact on gender and ethnicity. *Education & Treatment of Children, 25(4),* 415-437.

Shaffer, D., Fisher, P., Lucas, C.P., Dulcan, M.K. & Schwab-Stone, M.E. (2000). NIMH diagnostic interview schedule for children version iv (nimh disc-iv): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39(1),* 28-38.

Shaffer, D., Gould, M.S., Brasic, J., Ambrosini, P, Fisher, P., Bird, H. & Aluwahlia, S. (1983). A children's global assessment scale (cgas). *Archives of General Psychiatry, 40,* 1228-1231.

Sharma, S. & Kumar, A. (2006). Cluster analysis and factor analysis. In Grover, R. & Vriens, M. (Eds.), *The handbook of market research: Uses, misuses, and future advances* (365-393). Thousand Oaks, CA: Sage Publications.

Skovgaard, A.M., Houmann, T., Landorph, S.L. & Christiansen, E. (2004). Assessment and classification of psychopathology in epidemiological research of children 0-3 years of age: A review of the literature. *European Child & Adolescent Psychiatry, 3(6),* 337-346.

Smith, K.G. & Corkum, P. (2007). Systematic review of measures used to diagnose attention deficit/hyperactivity disorder in research on preschool children. *Topics in Early Childhood Special Education, 27(3),* 164-173.

Sonuga-Barke, E.J.S., Auerbach, J., Campbell, S.B., Daley, D. & Thompson, M. (2005). Varieties of preschool hyperactivity: Multiple pathways from risk to disorder. *Developmental Science, 8(2),* 141-150.

Speece, D.L. (1994-1995). Cluster analysis in perspective. *Exceptionality, 5(1),* 31-44.

Speltz, M.L., McClellan, J., DeKlyen, M. & Jones, K. (1999). Preschool boys with oppositional defiant disorder: Clinical presentation and diagnostic change. *Journal of the American Academy of Child & Adolescent Psychiatry, 38(7),* 838-845.

Sprafkin, J., Gadow, K.D. & Nolan, E.E. (2001). The utility of a dsm-iv-referenced screening instrument for attention-deficit/hyperactivity disorder. *Journal of Emotional and Behavioral Disorders, 9(3),* 182-191.

Tandon, M., Si, X. & Luby, J. (2011). Preschool onset attention-deficit/hyperactivity disorder: Course and predictors of stability over 24 months. *Journal of Child and Adolescent Psychopharmacology, 21(4),* 321-330.

Thorell, L.B. & Wählstedt, C. (2006). Executive functioning deficits in relation to symptoms of

adhd and/or odd in preschool children. *Infant and Child Development, 15(5)*, 503-518.

Van Horn, M.L., Atkins-Burnett, S., Karlin, E., Ramey, S.L. & Snyder, S. (2007). Parent ratings of children's social skills: Longitudinal psychometric analyses of the social skills rating system. *School Psychology Quarterly, 22(2*), 162-199.

Volpe, R.J., DiPerna, J.C., Hintze, J.M. & Shapiro, E.S. (2005). Observing students in classroom settings: A review of seven coding schemes. *School Psychology Review, 34(4)*, 454-474.

Waschbusch, D.A. (2002). A meta-analytic examination of comorbid hyperactive-impulsive-attention problems and conduct problems. *Psychological Bulletin, 128(1),* 118-150.

Wilens, T.E., Biederman, J., Brown, S., Monuteaux, M., Prince, J., Spencer, T.J. (2002). Patterns of psychopathology and dysfunction in clinically referred preschoolers. *Journal of Developmental and Behavioral Pediatrics, 23 (Suppl 1),* S31-S36.

Willcutt, E.G., Nigg, J.T., Pennington, B.F., Solanto, M.V., Rohde, L.A., Tannock, R., Loo, Sk>, Carlson, C.L., McBurnett, K. & Lahey, B.B. (2012). Validity of dsm-iv attention deficit/ hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology, 121 (4)*, 991-1010.

Zhu, S., Wang, D. & Li, T. (2010). Data clustering with size constraints. *Knowledge-Based Systems, 23*, 883-889.

Table 1

Participant Demographics

	N	Percent
Gender		
Male	52	74.3%
Female	18	25.7%
Ethnicity		
White/Non-Hispanic	51	72.9%
Black or African American	1	1.4%
Hispanic or Latino	8	11.4%
Other	10	14.3%
Age		
3	22	31.4%
4	32	45.7%
5	16	22.9%
Diagnosis (per screening measures)		
ADHD, Inattentive	2	2.86%
ADHD, Hyperactive-Impulsive	7	10.00%
ADHD, Combined	6	8.57%
Comorbid ADHD and ODD	55	78.57%

Table 2

Comparison of Demographic Characteristics of Current Sample versus Longitudinal Sample

Demographic Characteristic		Longitudinal Sample	Current Sample	Chi-Square	Sig.
Age (in years)				6.09	0.05
3	Count	9.0	22.0		
	Expected Count	14.9	16.1		
4	Count	35.0	32.0		
	Expected Count	32.3	34.7		
5	Count	21.0	16.0		
	Expected Count	17.8	19.2		
Gender				1.03	0.31
Male	Count	53.0	52.0		
	Expected Count	50.6	54.4		
Female	Count	12.0	18.0		
	Expected Count	14.4	15.6		
Ethnicity				2.15	0.54
White/Non-Hispanic	Count	42.0	51.0		
	Expected Count	43.7	49.3		
Black/African American	Count	3.0	1.0		
	Expected Count	1.9	2.1		
Hispanic/Latino	Count	10.0	8.0		
	Expected Count	8.5	9.5		
Maternal Education				3.02	0.56
Some High School	Count	3.0	4.0		
	Expected Count	3.2	3.8		
High School or GED	Count	16.0	14.0		
	Expected Count	13.6	16.4		
Some College or post High-School	Count	22.0	22.0		
	Expected Count	19.9	24.1		

College Graduate	Count	12.0	23.0		
	Expected Count	15.8	19.2		
Advanced Graduate or Professional Degree	Count	4.0	6.0		
	Expected Count	4.5	5.5		
Paternal Education				5.65	0.34
Eighth Grade or Less	Count	0.0	1.0		
	Expected Count	0.4	0.6		
Some High School	Count	5.0	7.0		
	Expected Count	5.2	6.8		
High School or GED	Count	16.0	22.0		
	Expected Count	16.6	21.4		
Some College or post High-School	Count	18.0	12.0		
	Expected Count	13.1	16.9		
College Graduate	Count	9.0	19.0		
	Expected Count	12.2	15.8		
Advanced Graduate or Professional Degree	Count	4.0	6.0		
	Expected Count	4.4	5.6		
Medication Status				1.05	0.31
No	Count	54.0	62.0		
	Expected Count	55.8	60.2		
Yes	Count	8.0	5.0		
	Expected Count	6.2	6.8		
Days per Week Attending an Educational Activity				6.14	0.29
0	Count	3.0	0.0		
	Expected Count	1.4	1.6		
2	Count	5.0	10.0		
	Expected Count	6.9	8.1		
3	Count	12.0	14.0		
	Expected Count	12.0	14.0		
4	Count	7.0	11.0		

Expected Count	8.3	9.7

5	Count	28.0	30.0
	Expected Count	26.8	31.2
6	Count	1.0	0.0
	Expected Count	0.5	0.5

Table 3

Average Inter-Observer Agreement for Abikoff Classroom Observation Code (COC) Observations

Behavior	Occurrence	Non-Occurrence	Kappa
Gross Motor Standing	81% (Range: 0% - 100%)	99% (Range: 95% - 100%)	0.82 (Range: 0.0 – 1.0)
Gross Motor Vigorous	83% (Range: 0% - 100%)	99% (Range: 88% - 100%)	0.85 (Range: 0.0 – 1.0)
Interference	70% (Range: 0% - 100%)	97% (Range: 85% - 100%)	0.75 (Range: 0.0 – 1.0)
Minor Motor Movement	75% (Range: 20% - 100%)	91% (Range: 52% - 100%)	0.78 (Range: 0.26 – 1.0)
Off-Task	81% (Range: 0% - 100%)	99% (Range: 89% - 100%)	0.84 (Range: 0.0 – 1.0)
Out Of Chair	85% (Range: 0% - 100%)	99% (Range: 88% - 100%)	0.87 (Range: 0.0 – 1.0)
Physical Aggression	99% (Range: 63% - 100%)	100% (Range: 96% - 100%)	0.99 (Range: 0.75 – 1.0)
Solicitation of Teacher	75% (Range: 0% - 100%)	99% (Range: 94% - 100%)	0.78 (Range: 0.0 – 1.0)
Teacher Approval	80% (Range: 0% - 100%)	99% (Range: 96% - 100%)	0.83 (Range: 0.0 – 1.0)
Threat/Verbal Aggression to Children	91% (Range: 0% - 100%)	100% (Range: 97% - 100%)	0.92 (Range: 0.0 – 1.0)
Threat/Verbal Aggression to Teacher	100% (Range: 100% - 100%)	100% (Range: 100% - 100%)	1.0 (Range: 1.0 – 1.0)

Table 4

Average Inter-Observer Agreement for Early Screening Project (ESP) Observations

Behavior	Occurrence	Non-Occurrence	Kappa
Solitary Play	81% (Range: 17% - 100%)	94% (Range: 65% - 100%)	0.84 (Range: 0.22 – 1.0)
Negative Physical	90% (Range: 0% - 100%)	100% (Range: 96% - 100%)	0.91 (Range: 0.0 – 1.0)
Parallel Play	84% (Range: 45% - 100%)	85% (Range: 38% - 100%)	0.81 (Range: 0.26 - 1.0)
Disruptive Behavior	83% (Range: 0% - 100%)	99% (Range: 86% - 100%)	0.84 (Range: -0.01 - 1.0)
Positive Social Engagement	82% (Range: 30% - 100%)	86% (Range: 34% - 100%)	0.81 (Range: 0.28 - 1.0)
Off-Task	81% (Range: 0% - 100%)	98% (Range: 79% - 100%)	0.84 (Range: 0.0 - 1.0)
Activity Change	71% (Range: 0% - 100%)	98% (Range: 92% - 100%)	0.76 (Range: -0.01 - 1.0)

Table 5

Summary of Correlations among All Variables Considered for Inclusion in Analysis

Strength of Correlation	Number of Statistically Significant Correlations between Variables
0.0 – 0.09	0
0.1 – 0.19	0
0.2 – 0.29	260
0.3 – 0.39	215
0.4 – 0.49	142
0.5 – 0.59	85
0.6 – 0.69	32
0.7 – 0.79	26
0.8 – 0.89	10
0.9 – 0.99	3

Table 6

Agglomeration Schedule for Hierarchical Cluster Analysis

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	7	8	1.740	0	0	29
2	18	19	12.094	0	0	23
3	13	32	23.365	0	0	6
4	37	68	35.251	0	0	8
5	38	40	47.541	0	0	11
6	13	54	60.598	3	0	25
7	20	27	74.475	0	0	28
8	10	37	89.391	0	4	12
9	45	58	104.462	0	0	21
10	9	22	120.343	0	0	31
11	38	49	136.747	5	0	36
12	10	46	153.508	8	0	26
13	16	70	170.422	0	0	21
14	15	51	187.353	0	0	30
15	52	64	204.406	0	0	37
16	33	59	222.556	0	0	45
17	4	34	241.011	0	0	39
18	1	31	259.698	0	0	41
19	39	44	278.692	0	0	34
20	25	65	298.760	0	0	30
21	16	45	318.985	13	9	46
22	26	67	339.216	0	0	24
23	18	43	359.868	2	0	28
24	24	26	381.307	0	22	39
25	13	66	402.894	6	0	34
26	10	28	424.652	12	0	36
27	50	55	447.509	0	0	45
28	18	20	470.545	23	7	47
29	5	7	494.745	0	1	40
30	15	25	519.591	14	20	38
31	9	11	544.958	10	0	38
32	63	69	570.730	0	0	43

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
33	2	47	597.190	0	0	57
34	13	39	624.090	25	19	50
35	6	36	651.451	0	0	48
36	10	38	679.087	26	11	43
37	42	52	707.164	0	15	47
38	9	15	738.005	31	30	42
39	4	24	769.308	17	24	50
40	5	23	800.652	29	0	56
41	1	62	834.495	18	0	59
42	9	48	868.888	38	0	55
43	10	63	903.790	36	32	46
44	21	61	939.771	0	0	49
45	33	50	975.994	16	27	49
46	10	16	1014.497	43	21	62
47	18	42	1057.176	28	37	53
48	6	41	1103.215	35	0	55
49	21	33	1150.509	44	45	60
50	4	13	1202.628	39	34	56
51	3	14	1255.060	0	0	63
52	53	60	1307.671	0	0	59
53	18	30	1361.161	47	0	57
54	12	56	1418.966	0	0	67
55	6	9	1476.986	48	42	65
56	4	5	1536.746	50	40	58
57	2	18	1599.768	33	53	65
58	4	17	1663.282	56	0	62
59	1	53	1730.633	41	52	61
60	21	29	1805.740	49	0	63
61	1	35	1888.531	59	0	68
62	4	10	1972.171	58	46	64
63	3	21	2057.704	51	60	66
64	4	57	2151.395	62	0	69
65	2	6	2257.101	57	55	66
66	2	3	2376.065	65	63	67
67	2	12	2501.634	66	54	68

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
68	1	2	2658.700	61	67	69
69	1	4	2967.000	68	64	0

Table 7

Comparison of Demographic Characteristics of Cluster 1 Versus Cluster 2

Demographic Characteristic		Cluster 1	Cluster 2	Chi-Square	Sig.
Age (in years)				0.67	0.72
3	Count	11.0	11.0		
	Expected Count	12.4	9.6		
4	Count	19.0	12.0		
	Expected Count	17.5	13.5		
5	Count	9.0	7.0		
	Expected Count	9.0	7.0		
Gender				0.42	0.52
Male	Count	30.0	21.0		
	Expected Count	28.8	22.2		
Female	Count	9.0	9.0		
	Expected Count	10.2	7.8		
Ethnicity				0.96	0.81
White/Non-Hispanic	Count	28.0	22.0		
	Expected Count	28.3	21.7		
Black/African American	Count	1.0	0.0		
	Expected Count	0.6	0.4		
Hispanic/Latino	Count	4.0	4.0		
	Expected Count	4.5	3.5		
Maternal Education				2.28	0.69

Some High School	Count	1.0	3.0
	Expected Count	2.2	1.8
High School or GED	Count	8.0	6.0
	Expected Count	7.8	6.2
Some College or post High-School	Count	13.0	8.0
	Expected Count	11.7	9.3

College Graduate	Count	12.0	11.0		
	Expected Count	12.9	10.1		
Advanced Graduate or Professional Degree	Count	4.0	2.0		
	Expected Count	3.4	2.6		
Paternal Education				9.21	0.10
Eighth Grade or Less	Count	0.0	1.0		
	Expected Count	0.6	0.4		
Some High School	Count	3.0	4.0		
	Expected Count	3.9	3.1		
High School or GED	Count	16.0	5.0		
	Expected Count	11.8	9.2		
Some College or post High-School	Count	6.0	6.0		
	Expected Count	6.7	5.3		
College Graduate	Count	11.0	8.0		
	Expected Count	10.7	8.3		
Advanced Graduate or Professional Degree	Count	1.0	5.0		
	Expected Count	3.4	2.6		
Medication Status				0.68	0.41
No	Count	36.0	25.0		
	Expected Count	35.1	25.9		
Yes	Count	2.0	3.0		
	Expected Count	2.9	2.1		
Days per Week Attending an Educational Activity				2.94	0.40
2	Count	6.0	4.0		
	Expected Count	5.6	4.4		
3	Count	5.0	8.0		
	Expected Count	7.3	5.7		
4	Count	8.0	3.0		
	Expected Count	6.2	4.8		
5	Count	17.0	13.0		
	Expected Count	16.9	13.1		

Table 8

Tests of Between-Subjects Effects for All Variables Included in Analysis

Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Variables that did not differ significantly between clusters						
TRF: Somatic Problems	0.25	1	0.25	0.25	0.62	0.00
ESP: Off-Task	0.26	1	0.26	0.26	0.61	0.00
Abikoff: Teacher Approval	0.33	1	0.33	0.32	0.57	0.00
ESP: Parallel Play	0.42	1	0.42	0.42	0.52	0.01
Abikoff: Minor Motor Movement	0.51	1	0.51	0.51	0.48	0.01
Abikoff: Solicitation of Teacher	0.52	1	0.52	0.51	0.48	0.01
ESP: Positive Social Engagement	0.80	1	0.80	0.80	0.38	0.01
SSRS-P: Social Skills	1.02	1	1.02	1.02	0.32	0.01
ESP: Activity Change	1.07	1	1.07	1.07	0.31	0.02
TRF: Somatic Complaints	1.41	1	1.41	1.42	0.24	0.02
Abikoff: Physical Aggression	1.76	1	1.76	1.78	0.19	0.03
Abikoff: Threat/Verbal Aggression to Teacher	2.12	1	2.12	2.15	0.15	0.03
ESP: Solitary Play	2.14	1	2.14	2.18	0.14	0.03
Abikoff: Gross Motor Vigorous	2.62	1	2.62	2.68	0.11	0.04
ESP: Disruptive Behavior	2.62	1	2.62	2.69	0.11	0.04
Abikoff: Interference	2.79	1	2.79	2.86	0.10	0.04
Abikoff: Gross Motor Standing	2.78	1	2.78	2.85	0.10	0.04
Abikoff: Out Of Chair	2.88	1	2.88	2.96	0.09	0.04
Conners-T: Cognitive Problems/Inattention	3.57	1	3.57	3.71	0.06	0.05
Abikoff: Threat/Verbal Aggression to Children	3.77	1	3.77	3.93	0.05	0.05
Abikoff: Off-Task	4.31	1	4.31	4.53	0.04	0.06
CBCL: Delinquent Behavior	5.66	1	5.66	6.07	0.02*	0.08
CBCL: Somatic Problems	5.06	1	5.06	5.38	0.02*	0.07
ESP: Negative Physical	5.27	1	5.27	5.63	0.02*	0.08
Conners-P: Cognitive Problems/Inattention	6.31	1	6.31	6.84	0.01*	0.09
Conners-P: Hyperactivity	6.38	1	6.38	6.92	0.01*	0.09
Conners-T: Anxious/Shy	5.86	1	5.86	6.31	0.01*	0.08
Conners-T: Perfectionism	6.34	1	6.34	6.88	0.01*	0.09
CBCL: Anxious/Depressed	8.53	1	8.53	9.60	0.00**	0.12
TRF: Withdrawal	10.15	1	10.15	11.72	0.00**	0.15

Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
TRF: Social Problems	22.40	1	22.40	32.69	0.00**	0.32
Conners-T: Social Problems	8.07	1	8.07	9.00	0.00**	0.12
Variables that differed significantly between clusters						
CBCL: Social Problems	11.35	1	11.35	13.39	0.00***	0.16
Conners-P: Oppositional	10.92	1	10.92	12.79	0.00***	0.16
Conners-P: Anxious/Shy	12.98	1	12.98	15.76	0.00***	0.19
Conners-P: Perfectionism	14.58	1	14.58	18.23	0.00***	0.21
Conners-P: Social Problems	14.89	1	14.89	18.71	0.00***	0.22
SSRS-P: Problem Behaviors	16.35	1	16.35	21.12	0.00***	0.24
TRF: Thought Problems	20.53	1	20.53	28.81	0.00***	0.30
TRF: Delinquent Behavior	24.15	1	24.15	36.62	0.00***	0.35
Conners-T: Hyperactivity	16.80	1	16.80	21.88	0.00***	0.24
SSRS-T: Social Skills	16.98	1	16.98	22.20	0.00***	0.25
SSRS-T: Problem Behaviors	20.78	1	20.78	29.31	0.00***	0.30

* $p < 0.05$
** $p < 0.01$
*** $p \leq 0.001$

Table 9

Mean Z-Score per Cluster across All Variables Included in Analysis

Variable	Cluster 1 Mean Z-Score	Cluster 2 Mean Z-Score
CBCL: Anxious/Depressed	0.31	-0.39
CBCL: Social Problems	0.36	-0.45
CBCL: Delinquent Behavior	0.25	-0.32
CBCL: Somatic Problems	0.24	-0.30
Conners-P: Oppositional	0.35	-0.44
Conners-P: Cognitive Problems/Inattention	0.27	-0.34
Conners-P: Hyperactivity	0.27	-0.34
Conners-P: Anxious/Shy	0.38	-0.48
Conners-P: Perfectionism	0.41	-0.51
Conners-P: Social Problems	0.41	-0.52
SSRS-P: Social Skills	-0.11	0.14
SSRS-P: Problem Behaviors	0.43	-0.54
TRF: Withdrawal	0.34	-0.43
TRF: Somatic Complaints	0.13	-0.16
TRF: Social Problems	0.50	-0.63
TRF: Thought Problems	0.48	-0.61
TRF: Delinquent Behavior	0.52	-0.66
TRF: Somatic Problems	-0.05	0.07
Conners-T: Cognitive Problems/Inattention	0.20	-0.25
Conners-T: Hyperactivity	0.44	-0.55
Conners-T: Anxious/Shy	0.26	-0.32
Conners-T: Perfectionism	0.27	-0.34
Conners-T: Social Problems	0.30	-0.38
SSRS-T: Social Skills	-0.44	0.55
SSRS-T: Problem Behaviors	0.49	-0.61
Abikoff: Interference	0.18	-0.22
Abikoff: Off-Task	0.22	-0.28
Abikoff: Minor Motor Movement	-0.08	0.10
Abikoff: Gross Motor Standing	0.18	-0.22
Abikoff: Out Of Chair	0.18	-0.23
Abikoff: Physical Aggression	0.14	-0.18
Abikoff: Threat/Verbal Aggression to Children	0.21	-0.26
Abikoff: Threat/Verbal Aggression to Teacher	0.16	-0.20

Variable	Cluster 1 Mean Z-Score	Cluster 2 Mean Z-Score
Abikoff: Gross Motor Vigorous	0.17	-0.22
Abikoff: Solicitation of Teacher	0.08	-0.10
Abikoff: Teacher Approval	0.06	-0.08
ESP: Negative Physical	0.24	-0.31
ESP: Disruptive Behavior	0.17	-0.22
ESP: Off-Task	-0.05	0.07
ESP: Activity Change	0.11	-0.14
ESP: Solitary Play	-0.16	0.20
ESP: Parallel Play	0.07	-0.09
ESP: Positive Social Engagement	0.10	-0.12

Table 10

Chi Square Analysis

Diagnostic Category		Cluster Membership 1	Cluster Membership 2	Chi-Square	Sig.
ADHD: Inattentive				0.27	0.87
No	Count	38.0	30.0		
	Expected Count	37.9	30.1		
	% within ADHD-I	55.9%	44.1%		
	% within Cluster Membership	97.4%	96.8%		
Yes	Count	1.0	1.0		
	Expected Count	1.1	0.9		
	% within ADHD-I	50.0%	50.0%		
	% within Cluster Membership	2.6%	3.2%		
ADHD: Hyperactive/Impulsive				0.78	0.38
No	Count	34.0	29.0		
	Expected Count	35.1	27.9		
	% within ADHD-H	54.0%	46.0%		
	% within Cluster Membership	87.2%	93.5%		
Yes	Count	5.0	2.0		
	Expected Count	3.9	3.1		
	% within ADHD-H	71.4%	28.6%		
	% within Cluster Membership	12.8%	6.5%		
ADHD: Combined				0.87	0.77
No	Count	36.0	28.0		
	Expected Count	35.7	28.3		
	% within ADHD-C	56.3%	43.8%		
	% within Cluster Membership	92.3%	90.3%		
Yes	Count	3.0	3.0		
	Expected Count	3.3	2.7		
	% within ADHD-C	50.0%	50.0%		
	% within Cluster Membership	7.7%	9.7%		

Diagnostic Category		Cluster Membership 1	Cluster Membership 2	Chi-Square	Sig.
Oppositional Defiant Disorder				N/A	N/A
No	Count	39.0	31.0		
	Expected Count	39.0	31.0		
	% within ODD	55.7%	44.3%		
	% within Cluster Membership	100.0%	100.0%		
Yes	Count	0.0	0.0		
	Expected Count	0.0	0.0		
	% within ODD	0.0%	0.0%		
	% within Cluster Membership	0.0%	0.0%		
Comorbid ADHD and ODD				0.14	0.71
No	Count	9.0	6.0		
	Expected Count	8.4	6.6		
	% within ODD	60.0%	40.0%		
	% within Cluster Membership	23.1%	19.4%		
Yes	Count	30.0	25.0		
	Expected Count	30.6	24.4		
	% within ODD	54.5%	45.5%		
	% within Cluster Membership	76.9%	80.6%		

Notes: Diagnostic status is mutually exclusive; for example, participants who met DISC-IV criteria for ADHD-Combined Type as well as ODD were included only in the "Comorbid ADHD and ODD" group. A Chi square value was not calculated for the ODD diagnostic category because this variable was considered a "constant," as no participants (N = 0) met DISC-IV criteria for this diagnosis alone, due to exclusion criteria.

Table 11

Examination of Differences between Parent and Teacher Ratings on Subscales Demonstrating Significant Between-Cluster Differences

***Scale:* Subscale Reflecting a Significant Between-Cluster Difference**	**Corresponding *Scale*: Subscale**	**Status**	
		Ratings Trended in Same Direction But Did Not Reach Statistical Significance	**Removed from Analysis to Prevent Multicollinearity**
Parent Rating Scales			
CBCL: Social Problems	*TRF:* Social Problems	X	
Conners-P: Social Problems	*Conners-T:* Social Problems	X	
Conners-P: Oppositional	*Conners-T:* Oppositional		X
Conners-P: Anxious/Shy	*Conners-T:* Anxious/Shy	X	
Conners-P: Perfectionism	*Conners-T:* Perfectionism	X	
Teacher Rating Scales			
SSRS-T: Social Skills	*SSRS-P:* Social Skills	X	
TRF: Delinquent Behavior	*CBCL:* Delinquent Behavior	X	
Conners-T: Hyperactivity	*Conners-P:* Hyperactivity	X	
TRF: Thought Problems	*CBCL:* Thought Problems		X

Notes: The remaining two scales on which statistically significant between-cluster differences occurred consisted of a corresponding pair of parent- and teacher-ratings: SSRS-P: Problem Behavior and SSRS-T: Problem Behavior. Both were consistent, reflecting more concerns (on average) among participants in Cluster 1 than those in Cluster 2.

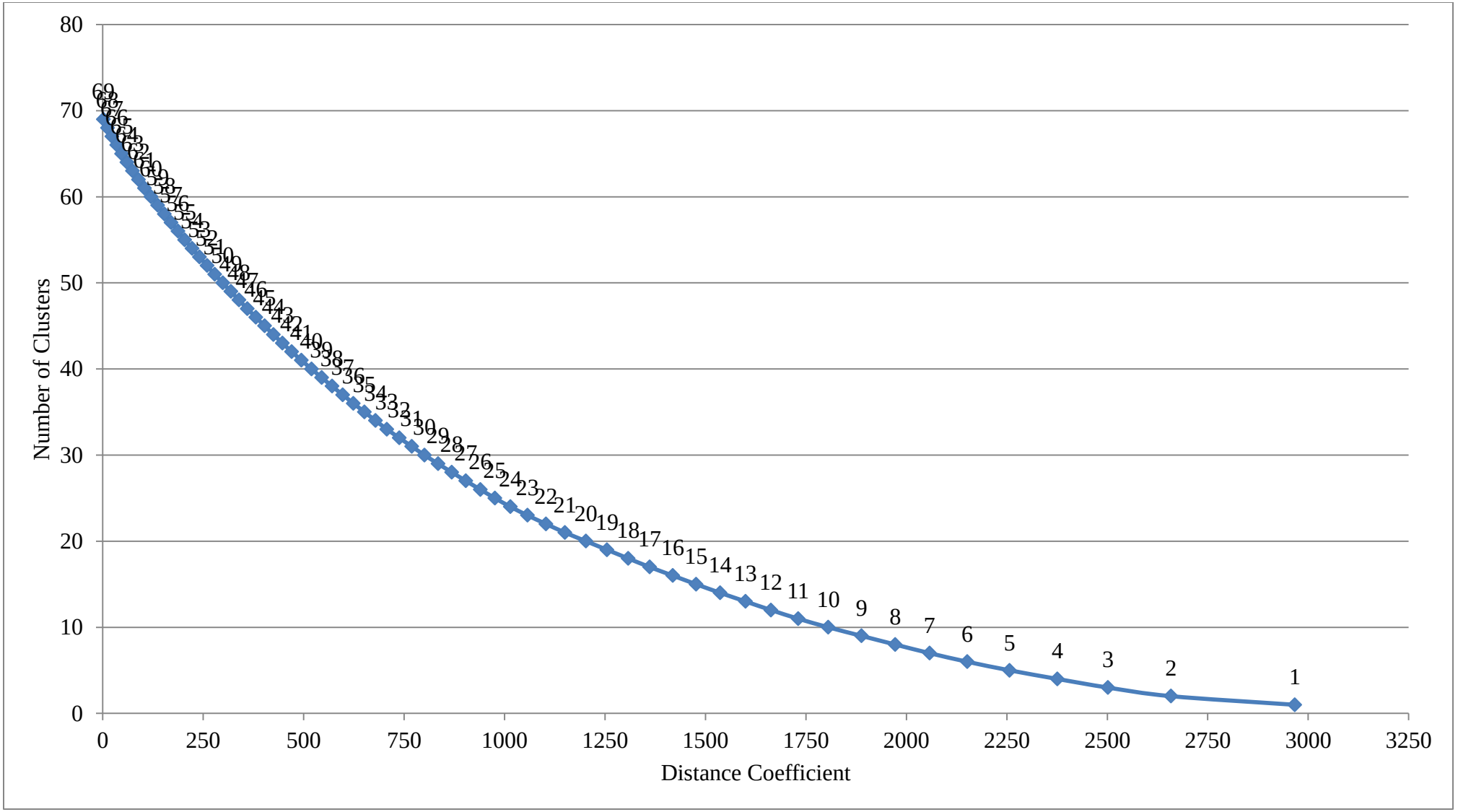

Figure 1. Distance coefficients for number of clusters in hierarchical cluster analysis

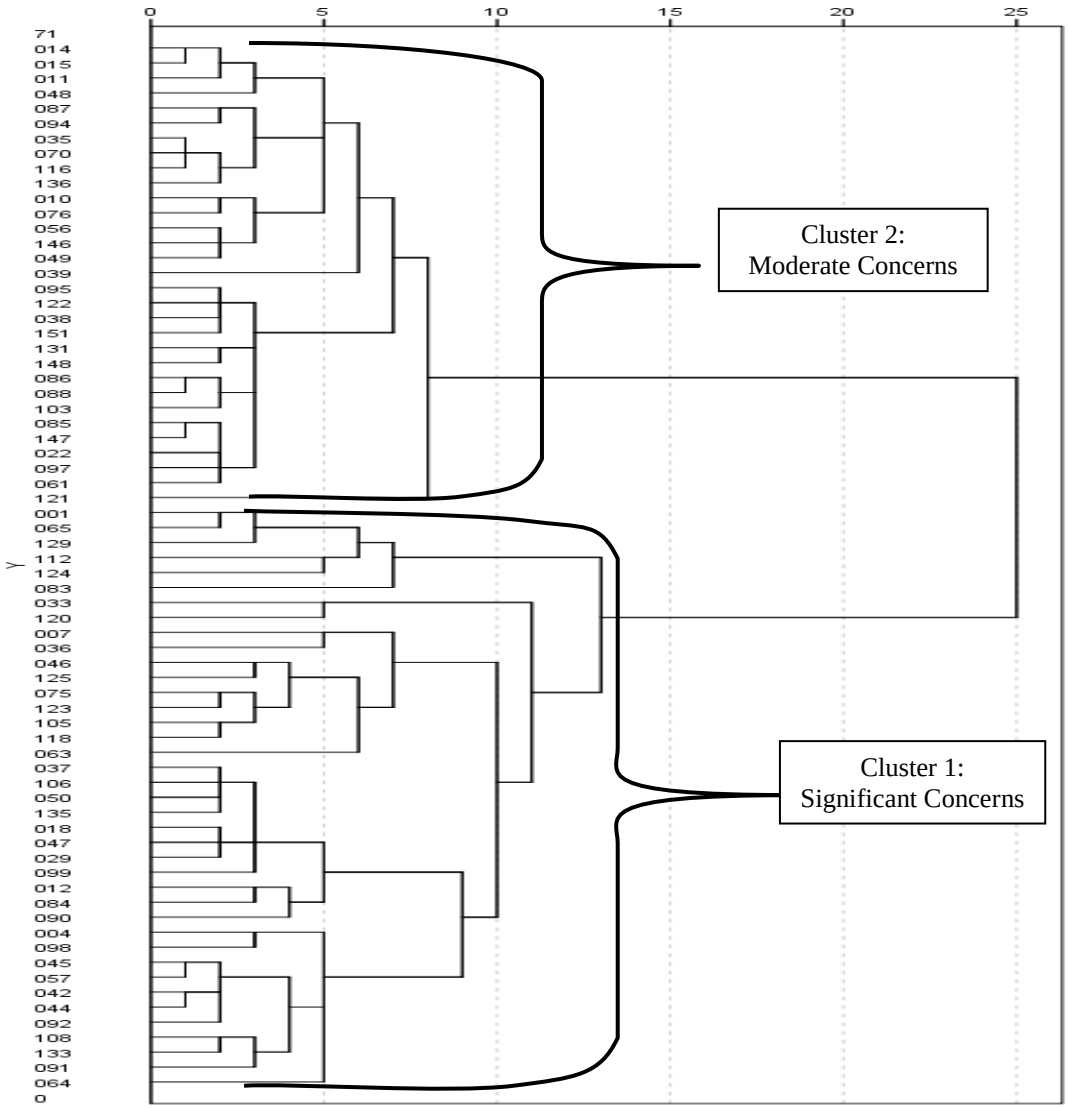

Figure 2. Dendrogram based on hierarchical cluster analysis

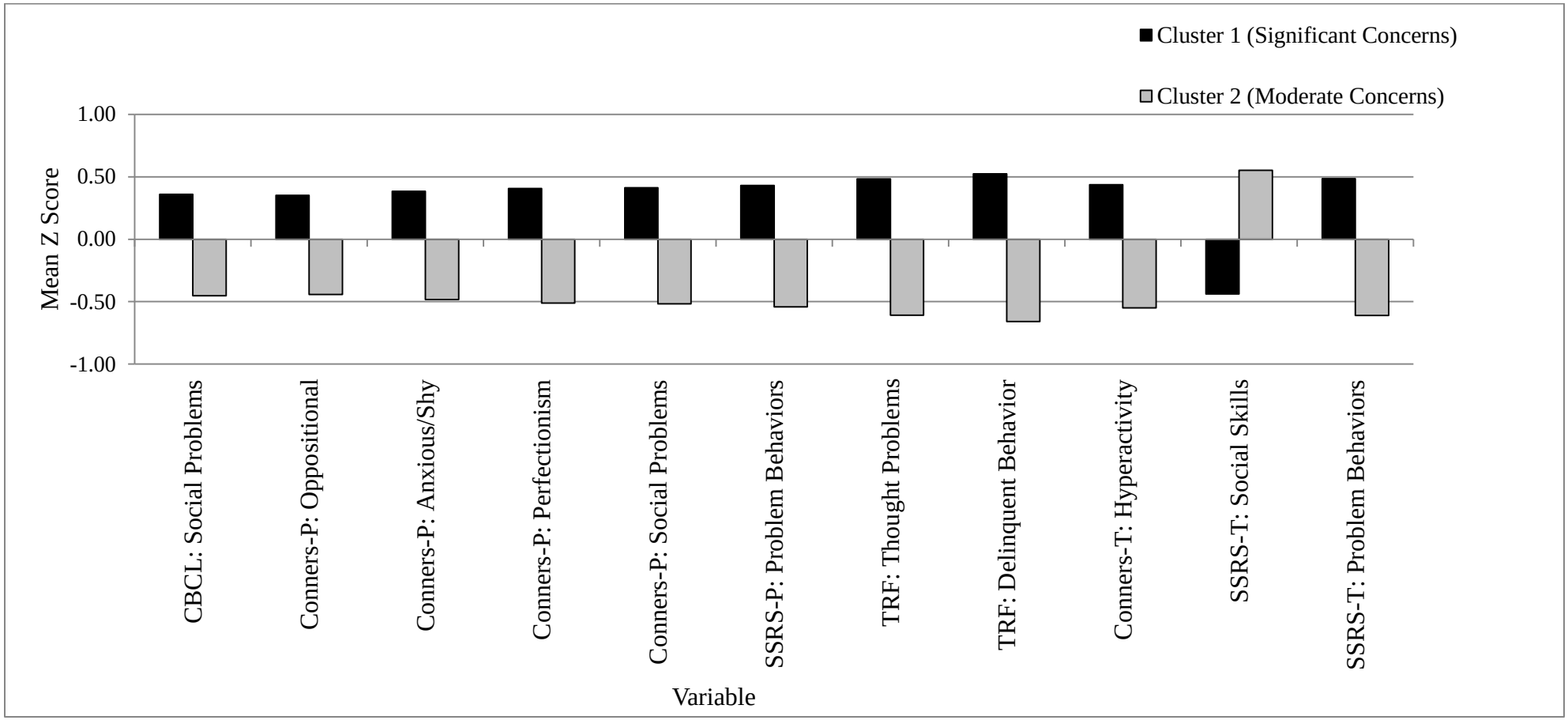

Figure 3. Mean Z-score by cluster

Note. Only variables for which there were significant between-cluster differences are depicted. Refer to Table 8 for a complete list of the mean cluster scores on each variable

Cluster 1: Significant Concerns

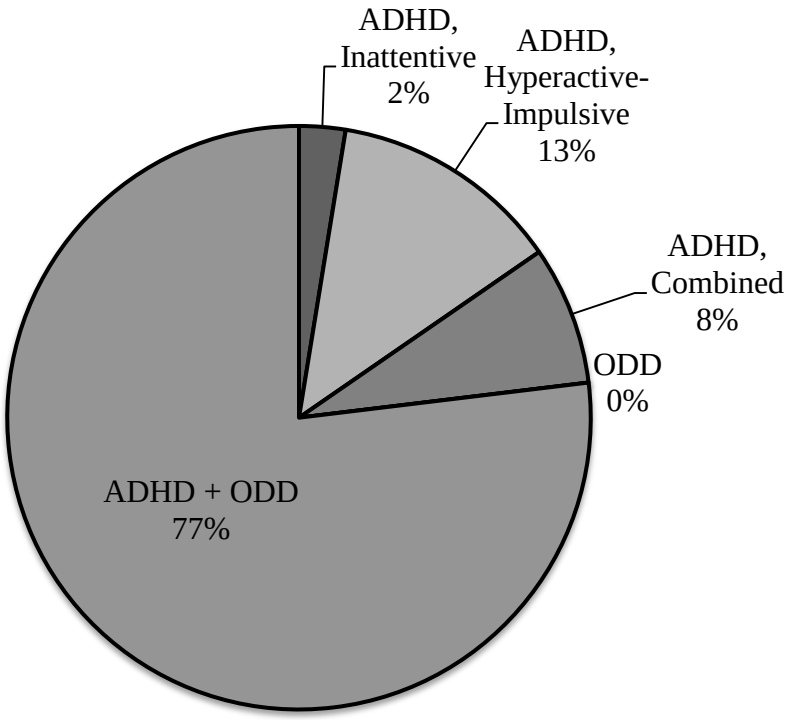

Figure 4. Diagnostic composition of cluster 1

Cluster 2: Moderate Concerns

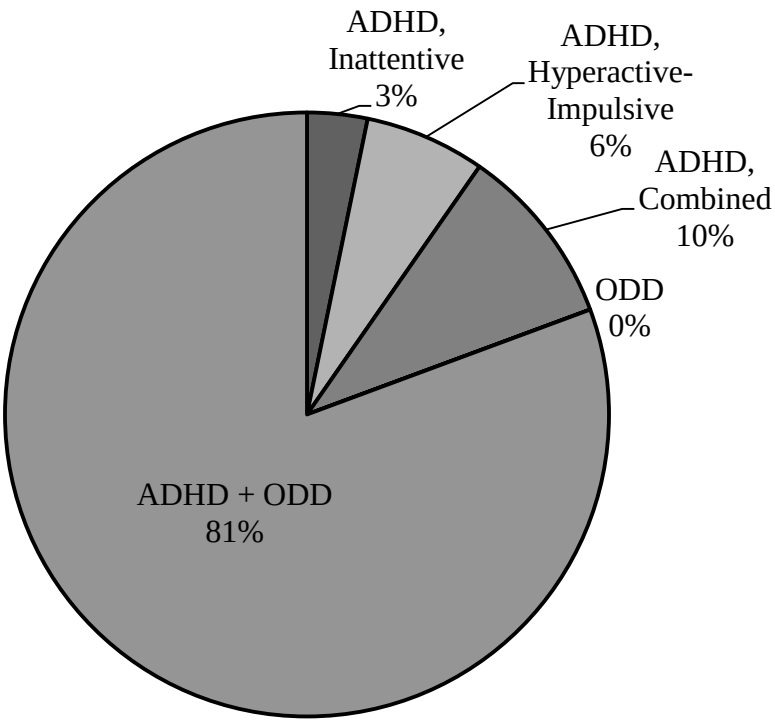

Figure 5. Diagnostic composition of cluster 2

Lisa M. DeKett

EDUCATION

Lehigh University
Bethlehem, Pennsylvania
Degree anticipated: Doctorate of Philosophy, School Psychology
Date anticipated: May 2014

Towson University
Towson, Maryland
Degree conferred: Master of Arts and Certificate of Advanced Study, School Psychology
May 2002

College of Notre Dame of Maryland
Baltimore, Maryland
Degree conferred: Bachelor of Arts Degree, Psychology
May 1997
Major in Psychology; Minor in English

CERTIFICATION

Nationally Certified School Psychologist (NCSP) October 2002 – present
Conferred by the National Association of School Psychologists (NASP)

Advanced Professional Certificate (APC) July 2008 – present
Conferred by Maryland State Department of Education (MSDE)

AWARDS/ACCOLADES

School Psychologist of the Year - Award Nominee (Frederick County Public Schools, May 2010)

School Psychology Graduate Student of the Year - Award Recipient (Towson University, May 2002)

Magna Cum Laude (College of Notre Dame of Maryland, May 1997)

SELECTED PROFESSIONAL EXPERIENCE

School Psychologist July 2008 – present
Frederick County Public Schools
Frederick, Maryland

School Psychology Intern August 2007 – June 2008
Howard County Public School System
Ellicott City, Maryland

Psychology Associate July 2007 – August 2007
Kennedy Krieger Institute
Baltimore, Maryland

Consultant September 2006 – June 2007
Project REACH
Lehigh University
Bethlehem, Pennsylvania

Graduate Assistant August 2004 – June 2006
Centennial Partial Hospitalization Program
Centennial School of Lehigh University
Bethlehem, Pennsylvania

Psychoeducational Evaluator January 2006 – May 2006
Fleetwood Area School District
Fleetwood, Pennsylvania

School Psychologist June 2002 – August 2004
C. Milton Wright High School
Harford County Public Schools
Bel Air, Maryland

School Psychology Intern August 2001 – June 2002
Charles County Public Schools
La Plata, Maryland

Graduate Assistant September 2000 – June 2001
School Psychology Department
Towson University, Maryland

PROFESSIONAL ACTIVITIES & EXPERIENCES

Presentations Provided

- Presented to school staff on procedures for completing a functional behavior assessment and developing a behavior intervention plan (Frederick County Public Schools, 2009)
- Presented to school psychology staff on best practices for working with English Language Learners (Frederick County Public Schools, 2009)
- Presented to Special Education Instructional Assistants about students with emotional disturbance, including definition/descriptions, behavioral manifestations, and interventions (Frederick County Public Schools, 2009)
- Co-presented at training workshops on the problem-solving process and the instructional intervention team model (Howard County Public Schools, 2007-2008)
- Presented to ESOL teachers on the research about retention and led a discussion about alternative ways of providing support to students (Howard County Public Schools, 2007-2008)

Participation on Teams and Committees

- PBIS Coach (Frederick County Public Schools, 2008 – present)
- Crisis Intervention Team (Frederick County Public Schools, 2009 – present; Howard County Public Schools, 2008 – 2009)
- Bilingual Assessment Committee member (Frederick County Public Schools, 2008 – 2009)
- Student Support Teams (Frederick County Public Schools, 2008 – present)
- Instructional Intervention Team at the elementary school level (Howard County Public Schools, 2007-2008)
- Kid-Talk teams at the middle school level (Howard County Public Schools, 2007-2008)
- Positive Behavior Intervention and Support teams at the elementary and middle school levels (Howard County Public Schools, 2007-2008)
- Culture Committee of the Psychological Services Staff (Howard County Public Schools, 2007-2008)
- Committee member: Best Practices in Risk Assessment for Potential Violence (Harford County Public Schools, 2004)

PROFESSIONAL AFFILIATIONS

National Association of School Psychologists 1999 – present
Member

Maryland School Psychologists' Association 1999 – present

Member

RESEARCH ACTIVITIES

Lehigh University January 2007 – May 2014
Dissertation research project examining the relationship between dimensional measures of behavior and diagnostic criteria for disruptive behavior disorders as applied to preschoolers.

Lehigh University September 2006 – June 2007
Consultant for Project REACH. Developing and implementing academic and behavioral interventions and monitoring the performance of students with emotional/behavioral disorders.

Lehigh University September 2005 - December 2006
Conducted research on the treatment practices of child and adolescent psychiatrists with regard to the use of multiple psychotropic medications.

Towson University October 2001 – December 2001
"Addressing the Function of Behavior through Understanding"
Research on educators' knowledge of the functions of various student behaviors and use of functional assessment in addressing behavioral concerns. Paper presented at the National Association of School Psychologists Chicago, IL. March 2002.

College of Notre Dame of Maryland **1997 – 1998**

Research on the correlation between various measures of focused and sustained attention and a diagnosis of Attention Deficit Hyperactivity Disorder. Presented at Eastern Psychological Association, April, 1997.

American Psychological Association – Ethics Board Fall 1997
Research on psychologists' views of the ethics of various types of clinician – client interactions/relationships. Presented at Eastern Psychological Association, April, 1997.